



(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 043 399 A2

(12)

EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
11.10.2000 Patentblatt 2000/41

(51) Int. Cl.⁷: **C12N 15/86**, C12N 7/01,
C12N 7/04, C12N 5/10,
C07K 14/18, A61K 49/00,
A61K 48/00

(21) Anmeldenummer: 00105929.4

(22) Anmeldetag: 23.03.2000

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: 03.04.1999 DE 19915178

(71) Anmelder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**

(72) Erfinder:
**Bartenschlager, Ralf, Dr.
55239 Gau-Odernheim (DE)**
(74) Vertreter: **Rudolph, Ulrike, Dr.
Patentanwältin
In der Schanz 10
69198 Schriesheim (DE)**

(54) Hepatitis C Virus Zellkultursystem

(57) Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepatoma-Zellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

Beschreibung

- [0001] Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d.h. die mit HCV-spezifischem Genmaterial transfiziert sind.
- [0002] Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Leberkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80-90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60 % aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.
- Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.
- [0003] Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50-60 nm und einer mittleren Dichte von 1,03-1,1g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989: Science, 244, 359-362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp.931-960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartenschlager 1997, Intervirology 40, 378 und vgl. Fig. 1 A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.
- Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozesierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.
- [0004] Das offene Leseraster ist an seinem 5' Ende von einer ca 340 Nukleotide langen nicht-translatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3' Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3' NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem amino-terminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.
- [0005] Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschriftsmäßig bei allen Blutkonserven angewendet werden (müssen).

- [0006] Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Komplettsequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z.T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.
- [0007] Die klonierten und charakterisierten Teil- und Komplettsequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, *Cell*, 87,343; Yem et al., 1998, *Protein Science*, 7, 837; Love et al., 1996, *Cell*, 87, 311; Kim et al., 1998, *Structure*, 6, 89; Yao et al., 1997, *Nature Structural Biology*, 4, 463, Cho et al., 1998, *J. Biol. Chem.*, 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.
- [0008] Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von 'rational drug design' als auch mit Hilfe von 'high throughput screens' weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiermodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.
- [0009] Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998, *J. Virol.* 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.
- [0010] Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können, (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences, USA*, 90, 6037-6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219-7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, *Journal of General Virology*, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Protein in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.
- [0011] Des Weiteren ist aus den Publikationen von Yoo et al. (1995, *Journal of Virology*, 69, 32-38) und von Dash et al., (1997, *American Journal of Pathology*, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels *in vitro* Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d.h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeföhrten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen. Außerdem ist im Stand der Technik mittlerweile bekannt (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), daß die hochkonservierte 3' NTR essentiell ist für die Virusreplikation, was in klarem Widerspruch zu den Behauptungen von Yoo et al. und Dash et al. steht, die für ihre Versuche in Unkenntnis des authentischen 3' Endes des HCV-Genoms ausschließlich HCV-Genome mit verkürzten 3' NTRs verwendet haben.
- [0012] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.
- [0013] Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Zellen, insbesondere Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das im

wesentlichen die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markerogen (Selektionsgen) umfaßt. "NTR" steht hier und im folgenden für "nicht-translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig. Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d.h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuBl 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

[0013] Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

[0014] Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markerogen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markerogen (Selektionsgen) vermittelt wird. Das Markerogen (Selektionsgen) scheint somit einerseits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

[0015] Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markerogen (Selektionsgen) umfaßt.

[0016] Der Begriff 5' NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3' NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

[0017] Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV - Replikation, - Pathogenesis und - Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann - als vollständiges Genom oder als Subgenom - gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

[0018] Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

[0019] Eine Ausführungsform des HCV-RNA-Konstrukt, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO:1 umfaßt. Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO:2 oder SEQ ID NO:3 oder SEQ ID NO:4 oder SEQ ID NO:5 oder SEQ ID NO:6 oder SEQ ID NO:7 oder SEQ ID NO:8 oder SEQ ID NO:9 oder SEQ ID NO:10 oder SEQ ID NO:11 umfassen.

[0020] Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3' NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

50

55

- 5 (a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTA
GCTTTTTTTTTCTTTTTGAGAGAGAGTCTCACTCTGTTGCC
AGACTGGAGT

10 (b) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTC TTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCCAGACT GGAGC

15 (c) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTAATCTT TTTTTTTCT TTTTTTTGA GAGAGAGAGT CTCACTCTGT
TGCCCAGACT GCAGC

20 (d) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTC TTTTTTTTT TCTTTTTTGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT

25 (e) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT TCTTTTTTGAGAGAGAG AGTCTCACTC
TGTTGCCAG ACTGGAGT

30 (f) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT TCTTTTTTGAGAGAGAG AGTCTCACTC
CTGTTGCCCA GACTGGAGT

35 (g) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT CTTTTTTTGAGAGAGAGA
GTCTCACTCT GTTGCCAGA CTGGAGT

40 (h) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAACT TTTTTTTTT TTTTCCTTT TTTTGAGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T

45 (i) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAACT TTTTTTTTT TTTTCCTTT TTTTGAGAGAG
AGAGAGTCTCA CTCTGTTGC CCAGACTGGAG GT

50 Das in den erfundungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Marker gen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z.B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter 'Antibiotikum' wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z.B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blasticidin.

[0021] Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

[0022] Eine Alternative zu den Antibiotikumresistenzgenen ist z.B. das Thymidin-Kinase-Gen, mit dem eine HAT-

Selektion durchgeführt werden kann.

[0023] Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5' NTR, d.h. strangabwärts der 5' NTR bzw. strangauwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion im Bereich der 3' NTR oder an anderer Stelle des HCV-Genoms oder -Subgenoms, z.B. innerhalb des Polyproteins.

[0024] Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

[0025] Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukt mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozys in die Zellen (z.B. mittels Transfektion eines Ribozymkonstrukt oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

[0026] Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

[0027] Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d.h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

[0028] Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

[0029] Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. ELISA) bestimmt werden.

[0030] Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder — allgemein — für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

[0031] Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z.B. Ubiquitin) oder über ein selbstspaltendes Peptid (z.B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z.B. in der Reihenfolge: Marker- bzw. Resistenzgen — interne Ribosomenbindungsstelle — Reportergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms einkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

[0032] Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z.B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z.B. die 5' NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z.B. auf Grund einer Reduktion der Translation des HCV-Polyproteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels 'rational drug design' oder 'high-throughput screening' am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.

- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflußt werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- 5 • Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapierezistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder - Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder 10 anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Rekonstruktion der HCV-RNA (z.B. mittels RT-PCR) und Sequenzanalyse können die für Therapierezistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausche/s in das Ursprungskonstrukt kann deren 15 Kausalität für die Therapierezistenz bewiesen werden.
- Die Produktion von authentischen Virusproteinen (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.
- 20 • Die Produktion von HCV Viren und virus-ähnlichen Partikeln insbesondere zur Entwicklung oder Herstellung von Therapeutika und Impfstoffen sowie für diagnostische Zwecke. Insbesondere zellkultur-adaptierte vollständige HCV-Genome, die mit dem erfindungsgemäßen Zellkultursystem hergestellt werden können, sind in der Lage, mit hoher Effizienz in Zellkulturen zu replizieren. Diese Genome besitzen alle Funktionen des HCV und sind deshalb 25 in der Lage infektiöse Viren zu produzieren.

25 [0033] Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:

- Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen: Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertionen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d.h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringelter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
- 30 • Die Konstruktion von HCV-RNA-Konstrukt mit integrierten Fremdgenen, beispielsweise zur Verwendung als leberzellspezifische Genfären in der Gentherapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese 35 dem Fachmann unter dem Begriff der 'Transkomplementation' bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
- 40 • Das Auffinden permissiver Zellen, d.h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das allerdings zunächst gemäß vorgenanntem Beispiel in einer Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumselektion unterzogen oder, 45 in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine Antibiotikumresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige 50 Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.

55 [0034] Das erfindungsgemäße Zellkultursystem erlaubt auch das gezielte Auffinden von HCV-RNA-Konstrukten, bei denen es auf Grund von Mutationen, die sich entweder zufällig im Rahmen der HCV-RNA-Replikation ereignen oder die gezielt in das Konstrukt eingeführt werden, zu einer Steigerung der Replikationseffizienz kommt. Solche Mutationen, die zu einer Veränderung der Replikation des HCV-RNA-Konstrukts führen, sind dem Fachmann als adaptive

Mutationen bekannt. Die Erfindung umfaßt deshalb auch Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts gemäß vorstehender Beschreibung, wobei die Mutanten gegenüber dem originären HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen. Sie umfaßt des Weiteren ein Verfahren zur Herstellung von Mutanten eines HCV-RNA-Vollängengenoms oder eines HCV-RNA-Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-RNA-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, sowie zellkultur-adaptierte Mutanten von HCV-RNA-Konstrukten, HCV-Vollängengenomen und HCV-Teilgenomen mit im Vergleich zu den ursprünglichen Konstrukten, Teil- oder Vollängengenomen erhöhter Replikationseffizienz.

[0035] Das erfindungsgemäße Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines erfindungsgemäßen HCV-RNA-Konstrukts, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, ist dadurch gekennzeichnet, daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 19 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

[0036] Bei einer vorteilhaften Weiterbildung dieses Herstellungsverfahrens werden die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 eingeschleust, das dabei erhaltenes Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone geerntet und daraus die HCV-RNA-Konstrukte isoliert.

Mit dieser Verfahrensvariante kann der Grad der adaptiven Mutationen und damit der Grad der Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten noch gesteigert werden.

[0037] Das erfindungsgemäße Verfahren zur Herstellung von Mutanten eines HCV - Vollängengenoms oder eines HCV- Teilgenoms oder eines beliebigen HCV-RNA-Konstrukts mit im Vergleich zu dem ursprünglichen HCV- Vollängengenom oder - Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz zeichnet sich dadurch aus, daß man mit Hilfe eines der beiden vorstehend genannten Herstellungsverfahren eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt, diese aus den Zellen isoliert, mit im Stand der Technik bekannten Methoden kloniert und sequenziert und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Mutationen bestimmt, und diese Mutationen dann entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, welche die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängen- oder -teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

Zum Nachweis bzw. zur Verifizierung derjenigen Mutationen, die tatsächlich eine Veränderung der Replikation und insbesondere eine Replikationssteigerung bewirken, kann ein Test durchgeführt werden, bei dem die bestimmten Nukleotid- und/oder Aminosäureaustausche in das ursprüngliche HCV-RNA-Konstrukt eingeführt und dieses wiederum in Zellkultur eingeschleust wird. Wenn die eingeführte Mutation tatsächlich zu einer Steigerung der Replikation führt, sollte im Fall eines HCV-RNA-Konstrukts mit selektierbarem Markogen die Zahl der resistenten Zellklone bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten Konstrukt. Im Fall eines Konstrukts mit einem Reportergen sollte die Aktivität bzw. Menge des Reporters bei dem künstlich mutierten Konstrukt deutlich höher sein als bei dem unbehandelten.

[0038] Die erfindungsgemäßen zellkultur-adaptierten HCV-RNA-Konstrukte mit hoher Replikationseffizienz sind dadurch gekennzeichnet, daß sie durch Nukleotid- und/oder Aminosäureaustausche von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19 ableitbar sind und daß sie mit einem der beiden vorstehend genannten Herstellungsverfahren erhältlich sind.

[0039] Diese zellkultur-adaptierten HCV-RNA-Konstrukte können dazu verwendet werden, beliebige HCV-RNA-Konstrukte oder HCV-Vollängen- oder Teilgenome mit erhöhter Replikationseffizienz herzustellen. Dabei können sowohl Konstrukte mit einem selektierbaren Resistenzgen als auch Konstrukte ohne ein solches bzw. mit einem nicht-selektierbaren Reportergen (z.B. Luziferase) hergestellt werden, denn aufgrund der sehr hohen Replikationseffizienz des zellkultur-adaptierten HCV-RNA-Konstrukts kann dessen Replikation auch in nicht-selektionierten Zellen nachgewiesen werden.

Die erfindungsgemäßen zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms oder eines HCV-Teilgenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, sind dadurch charakterisiert, daß sie mit einem Verfahren erhältlich sind, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 19, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

[0040] Eine Gruppe ganz bevorzugter HCV-RNA-Konstrukte, HCV-Vollängengenome und HCV-Teilgenome mit hoher und sehr hoher Replikationseffizienz und infolgedessen sehr guter Eignung für die praktische Anwendung ist

EP 1 043 399 A2

dadurch gekennzeichnet, daß sie einen oder mehrere oder alle der in Tabelle 3 aufgelisteten Aminosäure- bzw. Nukleotidaustausche und/oder einen oder mehrere der folgenden Aminosäureaustausche aufweist: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile -> val. (Die Zahlen beziehen sich auf die Aminosäurepositionen des Polyproteins des HCV-Isolats con1, siehe Tabelle 1).

5

Besondere Eigenschaften der in den Sequenzprotokollen angegebenen Sequenzen:

SEQ ID-NO: 1

10 [0041]

Name: I389/Core-3'/wt

Aufbau (Nukleotidpositionen):

- 15 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 20 4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
- 25 9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
- 30 14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 2

35 [0042]

Name: I337/NS2-3'/wt

Aufbau (Nukleotidpositionen):

- 40 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
- 45 4. 1801-8403: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1801-2451: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
6. 2452-4344: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4345-4506: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
- 50 8. 4507-5289: Nichtstrukturprotein 4B (NS4B)
9. 5290-6630: Nichtstrukturprotein 5A (NS5A)
10. 6631-8403: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8407-8637: HCV 3' nicht-translatierte Region

55

SEQ ID-NO: 3

[0043]

5 Name: I389/NS3-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
- 15 5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
10. 7771-8001: HCV 3' nicht-translatierte Region

20 SEQ ID-NO: 4

[0044]

25 Name: I337/NS3-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 30 3. 1190-1800: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1801-7758: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B
5. 1801-3696: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
- 35 6. 3697-3858: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3859-4641: Nichtstrukturprotein 4B (NS4B)
8. 4642-5982: Nichtstrukturprotein 5A (NS5A)
9. 5983-7755: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
10. 7759-7989: HCV 3' nicht-translatierte Region

40 SEQ ID-NO: 5

[0045]

45 Name: I389/NS2-3'/wt

Aufbau (Nukleotidpositionen):

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
- 50 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-8418: HCV Polyprotein von Nichtstrukturprotein 2 bis Nichtstrukturprotein 5B
5. 1813-2463: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
- 55 6. 2464-4356: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
7. 4357-4518: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
8. 4519-5301: Nichtstrukturprotein 4B (NS4B)
9. 5302-6642: Nichtstrukturprotein 5A (NS5A)

EP 1 043 399 A2

10. 6643-8415: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
11. 8419-8649: HCV 3' nicht-translatierte Region

SEQ ID-NO: 6

5

[0046]

Name: I389/NS3-3'/9-13F
Aufbau (Nukleotidpositionen):

10

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

15

20

SEQ ID-NO: 7

[0047]

Name: I389/Core-3'/9-13F
Aufbau (Nukleotidpositionen):

30

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 9-13F
5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

35

40

45

SEQ ID-NO: 8

[0048]

Name: I389/NS3-3'/5.1
Aufbau (Nukleotidpositionen):

50

1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker

EP 1 043 399 A2

3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1

5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5B (NS5A)
10. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 9

15 [0049]

Name: I389/Core-3'/5.1

Aufbau (Nukleotidpositionen):

20. 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 5.1
25. 5. 1813-2385: HCV Core Protein; Strukturprotein
6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
8. 4051-4239: Protein p7
30. 9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
35. 14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
15. 10846-11076: HCV 3' nicht-translatierte Region

SEQ ID-NO: 10

40 [0050]

Name: I389/NS3-3'/19

Aufbau (Nukleotidpositionen):

45. 1. 1-341: HCV 5' nicht-translatierte Region
2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
4. 1813-7767: HCV Polyprotein von Nichtstrukturprotein 3 bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
50. 5. 1813-3708: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
6. 3709-3870: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
7. 3871-4653: Nichtstrukturprotein 4B (NS4B)
8. 4654-5994: Nichtstrukturprotein 5A (NS5A)
55. 9. 5995-7767: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase 7771-8001: HCV 3' nicht-translatierte Region

SEQ ID-NO: 11

[0051]

- 5 Name: I389/Core-3'/19
 Aufbau (Nukleotidpositionen):
1. 1-341: HCV 5' nicht-translatierte Region
 2. 342-1193: HCV Core Protein-Neomycin Phosphotransferase Fusionsprotein; selektionierbarer Marker
 - 10 3. 1202-1812: Interne Ribosomenbindungsstelle des Encephalomyokarditis Virus; erlaubt die Translation des dahinterliegenden HCV offenen Leserasters
 4. 1813-10842: HCV Polyprotein von Core bis Nichtstrukturprotein 5B der zellkultur-adaptierten Mutante 19
 - 15 5. 1813-2385: HCV Core Protein; Strukturprotein
 6. 2386-2961: Hüllprotein 1 (envelope protein 1); Strukturprotein
 7. 2962-4050: Hüllprotein 2 (envelope protein 2); Strukturprotein
 8. 4051-4239: Protein p7
 9. 4240-4890: Nichtstrukturprotein 2 (NS2); HCV NS2-3 Protease
 10. 4891-6783: Nichtstrukturprotein 3 (NS3); HCV NS3 Protease/Helikase
 - 20 11. 6784-6945: Nichtstrukturprotein 4A (NS4A); NS3 Protease Kofaktor
 12. 6946-7728: Nichtstrukturprotein 4B (NS4B)
 13. 7729-9069: Nichtstrukturprotein 5A (NS5A)
 14. 9070-10842: Nichtstrukturprotein 5B (NS5B); RNA-abhängige RNA-Polymerase
 15. 10846-11076: HCV 3' nicht-translatierte Region

25 [0052] Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Tabellen und Figuren näher erläutert. Die erwähnten Figuren zeigen

- Fig. 1 A:** Die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5' NTR und 3' NTR) — als Horizontalbalken dargestellt —, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der 'GDD-katalytischen Domäne' der NS5B RNA Polymerase (GDD) und der Position der 3' Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) — oberhalb des Genomschemas eingezeichnet —. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.
 Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (Neo^R), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3' Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck (Δ) markiert.
- Fig. 1 B:** Das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen.
 Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben.
- Fig. 1 C** : Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone.
 Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle.
- Fig. 2 A:** Das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I₃₇₇/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13). Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I₃₇₇/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2 - 6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 µg 9-13 DNA vor der PCR (zwecks Ausschluß

eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem µg Gesamt-DNA des Zellkons 9-13 erhalten wurde.

- 5 **Fig. 2 B:** Das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA. Die Pfeile markieren die Positionen von Replikon-RNA. Die "plus" und "minus"-Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrand" und "Plusstrand" bezeichnen die Spezifität der radioaktiven RNA-Sonden.
- 10 **Fig. 2 C:** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin.
- 15 **Fig. 3 A:** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung. Die Spuren 7 - 9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben.
- 20 **Fig. 3 B:** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen.
- 25 **Fig. 4:** Schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5' HCV-IRES, dem Neomycinphosphotransferasegen (NeoR), einem heterologen IRES-Element, z.B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3' NTR.
- 30 **Fig. 5:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Antibiotikumresistenzgen (A) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3' NTR (bicistronische RNA).
- 35 **Fig. 6:** Schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit insertiertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; — das Reporterprotein wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z.B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schräffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, (D) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; — das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt).
- 40 **Fig. 7:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV 5' und 3' NTRs verbunden ist. Die dicken Linien stellen die HCV 5' und 3' NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar.
- 45 **Fig. 8:** Schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.
- 50 **Fig. 9:** Methodisches Vorgehen zum Vergleich der spezifischen Infektiosität (ausgedrückt als Anzahl gebildeter Zellkolonien) von Gesamt-RNA versus in vitro Transkripte. HCV-RNA wird mittels in vitro Transkription eines entsprechenden RNA-Konstrukts hergestellt und durch Messung der optischen Dichte bei 260 nm (OD 260 nm) quantifiziert. Eine definierte Anzahl dieser Moleküle wird mit einer bestimmten Menge Gesamt-RNA von naiven Huh-7 Zellen gemischt und diese Mischung mit Hilfe der Elektroporation in

naïve Huh-7 Zellen eingeschleust. Parallel dazu wird die Gesamt-RNA eines Zellklons, der mit der in Figur 1 beschriebenen Methode hergestellt wurde, mit einem im Stand der Technik bekannten Verfahren isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot unter Verwendung einer HCV-spezifischen RNA-Sonde und anschließender Quantifizierung mittels Phosphoimager bestimmt. Eine definierte Menge dieser Gesamt-RNA wird analog den in vitro Transkripten in naïve Huh-7 Zellen transfiziert. Diese Zellen in beiden Ansätzen werden danach einer G418-Selektion unterzogen und die Anzahl der gebildeten Kolonien durch Auszählen nach fixieren und anfärben mit Coomassie-Brilliant-Blau bestimmt. Zur Bestimmung der Transfektionseffizienz wird jedem Transfektionsansatz 1 µg eines Plasmids zugesetzt, das die Expression der Luziferase erlaubt. Ein Aliquot der transfizierten Zellen wird nach 24 Stunden geerntet und die Luziferaseaktivität im jeweiligen Zellysat bestimmt. Die Anzahl der Kolonien wird jeweils auf die Luziferaseexpression normiert.

- Fig. 10:** Sequenzanalyse der 9-13 Klone. Gesamt-RNA des Zellklons 9-13, der durch Transfektion des HCV-RNA-Konstrukt I₃₇₇/NS3-3' entstand, wurde mit einem im Stand der Technik bekannten Verfahren isoliert und das HCV-RNA-Konstrukt von Nukleotidposition 59 bis 9386 mit Hilfe der 'long-distance RT-PCR' unter Verwendung der primer S59 und A9413 amplifiziert. Die PCR-Fragmente wurden kloniert und 11 Klone (genannt 9-13 A - K) vollständig sequenziert, wobei sich die Klone D und I, E und G sowie H und J als identisch erwiesen. Die Positionen der Aminosäureunterschiede in der NS3-5B Region zwischen den reklierten HCV-RNAs und dem parentalen Konstrukt sind mit einem dicken vertikalen Strich beim jeweiligen Klon markiert. Jeder Klon wurde mit dem Restriktionsenzym *Sfi* 1 verdaut und das jeweilige Fragment in das parentale Konstrukt inseriert. Diese Klone wurden jeweils in Huh-7 Zellen transfiziert und die Zellen wie in Figur 1 beschrieben einer Selektion unterzogen. Die Anzahl der mit jedem Konstrukt erhaltenen Zellklone ist rechts neben dem jeweiligen Konstrukt vermerkt.
- Fig. 11 A:** Prinzip der Replikationsbestimmung mit Hilfe eines Reportergens. Im oberen Teil der Figur ist das HCV-DNA-Konstrukt I₃₈₉/Luc/NS3-3' dargestellt, bestehend aus der HCV 5' NTR (Nukleotidposition 1-389), dem Luziferasegenen (*luc*), der IRES des Encephalomyocarditis Virus, dem HCV NS3-5B und der 3' NTR. Die Position des aktiven Zentrums der NS5B RNA-Polymerase, in das ein inaktivierender Aminosäureaustausch eingeführt wurde, ist mit 'GND' angedeutet. Die Plasmide, die für das replikationskompetente bzw. das defekte HCV-RNA-Konstrukt kodieren, werden mit dem Restriktionsenzym *Sca* I verdaut und in eine in vitro Transkription mit der T7 RNA-Polymerase eingesetzt. Nach Entfernung der Matrizen-DNA werden die jeweiligen HCV-RNA-Konstrukte mittels Elektroporation in naive Huh-7 Zellen eingeschleust und diese in regelmäßigen Abständen geerntet.
- Fig. 11 B:** Vergleich der Luziferaseaktivitäten in Zellen transfiziert mit dem parentalen HCV-RNA-Konstrukt I₃₈₉/Luc/NS3-3'/wt (wt) oder den folgenden Varianten: Der inaktiven RNA (318 DN), der Variante 9-13F oder der Variante 5.1. Die Zellen wurden 6 (nicht gezeigt), 24, 48, 72, 96, 120, 144 und 168 Stunden nach der Transfektion geerntet und die Luziferaseaktivitäten luminometrisch bestimmt.
- Fig. 12:** Selektionierbare HCV-Vollängengenome (Konstrukte I₃₈₉/core-3'/5.1 und I₃₈₉/core-3'/9-13F).
- (A) Schematische Darstellung des Vollängenkonstrukt. Der Bereich zwischen den beiden ange deuteten Erkennungsstellen für das Restriktionsenzym *Sfi* I entspricht den Sequenzen der hoch adaptierten RNA-Varianten 5.1. oder 9-13F.
- (B) Anzahl der Kolonien die nach Transfektion von jeweils 0,1 µg in vitro transkribierter RNA der unter A dargestellten Konstrukte I₃₈₉/core-3'/5.1 in HUH7-Zellen erhalten wurden. Angegeben ist das Ergebnis eines repräsentativen Experiments.
- (C) Nachweis autonom replizierender HCV-Vollängen-RNAs in G418-resistenten Zellklonen, die nach Transfektion des entsprechenden in vitro Transkripts erhalten wurden. Die Abbildung zeigt das Autoradiogramm eines Northern Blots, der mit einer Sonde gegen das *neo*-Resistenzgen und der HCV 5' NTR hybridisiert wurde. Die in Spur 1 und 2 dargestellten Kontrollen entsprechen jeweils 10⁸ Molekülen der angegebenen in vitro Transkripte, gemischt mit Gesamt-RNA aus naiven Huh-7 Zellen. Die Negativkontrolle enthält ausschließlich Gesamt RNA aus naiven Huh-7 Zellen (Spur 3). Die Spuren 4-9 enthalten 3-10 µg Gesamt-RNA aus G418-resistenten Zellklonen, die nach Transfektion von in vitro transkribierter I₃₈₉/core-3'/5.1-RNA bzw. I₃₈₉/core-3'/9-13F-RNA erhalten wurden. Die für die Selektion verwendete G418-Konzentration ist jeweils angegeben. Fünf der dargestellten Zellklone enthalten die hoch adaptierte RNA-Variante 5.1 (Spur 4-8), einer die adaptierte RNA-Variante 9-13F (Spur 9).

- Fig. 13: HCV-RNA-Konstrukte mit einem Reportergen. (A) Bicistronische HCV-RNA-Konstrukte. Das Reportergen wird mit Hilfe einer separaten IRES translatiert. (B) Monocistronische HCV-RNA-Konstrukte. Das Reportergenprodukt wird als Fusionsprotein mit einem HCV-Protein exprimiert. Die beiden Anteile sind über eine Erkennungssequenz für eine virale oder zelluläre Protease verbunden, die eine proteolytische Trennung der beiden fusionierten Proteinanteile erlaubt. Im gezeigten Beispiel wurden das Reportergenprodukt und das jeweilige HCV-Protein über eine Erkennungssequenz für Ubiquitin (Ub) fusioniert.
- Fig. 14: Tricistronisches Vollängen HCV-RNA-Konstrukt, das zusätzlich zum Resistenzgen ein Fremdgen inseriert besitzt.
- Fig. 15: Monocistronische HCV-RNA-Konstrukte, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird. Das Resistenzgen (RG) ist entweder als Fusionsprotein aktiv oder es wird so mit einer proteolytisch spaltbaren Sequenz mit dem HCV-Anteil fusioniert, daß das Resistenzgenprodukt durch eine zelluläre oder virale Protease vom HCV-Anteil abgespalten wird. Im gezeigten Beispiel wurde das Resistenzgen über die für Ubiquitin (Ub) kodierende Sequenz mit dem jeweiligen HCV-Anteil fusioniert.

Beispiel 1: Herstellung von HCV-RNA-Konstrukten

20 (A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT- PCR

[0053] Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d.h. die HCV-RNA wie nachfolgend beschrieben isoliert:
[0054] Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacci (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103 (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGCCATTGG CCTGGAG) und dem 'expand reverse transcriptase'- System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR=polymerase chain reaction) durchgeführt, und zwar unter Verwendung des 'expand long template'-Systems (Boehringer Mannheim, Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und S4542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCACAGCCCGCGTCATGCACTCG) oder S4542 und A9386 (TTAGCTCCCCG TTCATCGGTTGG) eingesetzt. Nach 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (=3' NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten 'X-tails' (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (=5' NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5' Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor insertiert. Abweichungen von der Konsensussequenz wurden mittels ortsgerichteter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3' Ende herzustellen, wurde die 3'-NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon 'WS' nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3' Ende der 3' NTR (Kolykhalov et al. ibid.) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3' NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch

gefunden.

[0055] Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

(B) Synthese selektierbarer HCV-Subgenom-Konstrukte

[0056] Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, Molecularcloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5' NTR insertiert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1 A). Ebensogut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstruktinsertiert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5 A) oder in die 3' NTR (siehe Fig. 5 B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Encephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3' Ende des Genoms steuert.

[0057] Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt: Auf der Basis von Deletionsanalysen der 3' Grenze der HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) wurden verschiedene Abschnitte der 5' NTR mit dem NPT Gen fusioniert und anhand von Kotransfektionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1 A).

[0058] Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

[0059] Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I₃₇₇/NS2-3' (oder I₃₇₇/NS3-3') und I₃₈₉/NS2-3' (oder I₃₈₉/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

[0060] Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parental dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, *EMBO J.* 15, 12; und Lohmann et al., 1997, *J. Virol.* 71, 8416).

(C) Synthese selektierbarer HCV-Genom-Konstrukte

[0061] Ein NS2-3' Subgenomkonstrukt, das am 5' Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit Ncol und Spel restringiert und mittels präparativer Agarosegelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem Ncol/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/Spel-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3' NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit Pmel und Spel restringiert und mit dem analog restringierten I₃₈₉/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

(D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Trnaskripten

[0062] Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit Scal linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Trnkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7.5, 12,5 mM MgCl₂, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/ μ l, 50 μ g/ml restringierte DNA und ca. 2 Einheiten/ μ l T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J.C. Porte, 1991, Bio Techniques 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro μ g DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittel optischer Dichtemessungen quantifiziert und

ihrer Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

Beispiel 2: Transfektionsexperimente mit der Hepatomazelllinie Huh-7

- 5 [0063] Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1 D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol, saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.
- 10 [0064] Es wurden drei separate Transfektionsxperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, *Cancer Res.* 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8 x 10⁶ Huh-7-Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3 - 5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.
- 15 [0065] Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transzisiert worden waren und 1 Klon (Klone 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3' HCV-RNA transzisiert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.
- 20 [0066] Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermittelten könnte.
- 25 [0067] Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, *Anal. Biochem.* 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von 0,5 - 1 x 10⁶ Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1 B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 10⁹ in-vitro-Transkripte (ivtr.), die zu den I₃₈₉/NS2-3'/wt oder den I₃₈₉/NS3-3'/wt Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1 B dargestellt.
- 30 [0068] Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transzisierten Klons 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3' Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genomkonstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40 µg/ml, 1h, 37°C) in 10 mM Tris, pH 7,5, 1 mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend 4 - 8 x 10⁴ Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTTGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10⁷ Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transzisiert war, durchgeführt, und als Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt.
- 40 [0069] Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrolle.
- 45 [0070] Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrolle.
- 50 [0071] Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrolle.
- 55 [0072] Die Ergebnisse dieser Untersuchung sind in Fig. 1 C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrolle.

len, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3' Replikon/NS3-3' HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

[0069] Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1 B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in ~ 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren: Es wurden jeweils 10^6 - 10^2 Plasmid Moleküle ($I_{377}/NS3-3'/wt$) entweder direkt (Spuren 7 - 11) oder nach Zugabe von jeweils 1 µg 9-13 DNA (Spuren 2 - 6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem µg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder 2 µg Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ('plusstrand', obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ('minusstrand', untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2 B dargestellt.

Im Fall des Plusstrangs wurden ca. 10^8 Kopien/µg Gesamt-RNA nachgewiesen, was 1000 - 5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

[35] Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [3 H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [3 H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5×10^5 Zellen der Klone 9-13 und 8-1 mit 100 µ Ci [3 H]Uridin für 16 Std. in Abwesenheit (-) oder Gegenwart (+) von 4 µg/ml Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

[45] Die Ergebnisse dieser Analyse sind in Fig. 2 C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, EMBOJ. 15, 12 und Lohmann et al., 1997, J Virol. 71, 8416) war die Replikation der HCV RNA nicht durch Dactinomycin beeinflußt worden, während die Synthese von zellulärer RNA gehemmt worden war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Rekonstruktion der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der rekonstruierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transfizierten Transkript des HCV-Konstruktks $I_{377}/NS3-3'/wt$ übereinstimmt.

[0070] Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [35 S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z.B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nicht-denaturierenden Bedingungen (z.B. nach Bartenschlager et al., 1995, J. Virol. 69, 7519) und unter Verwendung von drei verschiedenen Antisera (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abge-

trennt worden.. Die Immunokomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonkonstrukt I₃₇₇/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7 - 9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum ('3/4') bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

[0071] Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transakter Expression desselben bicistronischen HCV-RNA-Konstrukturts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunfluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antiseren durchgeführt (z.B. nach Bartenschlager et al., 1995, *J. Virol.* 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azetin fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antiseren inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchens-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff 'Evans Blue' gegenfärbt.

[0072] Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antiseren war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antiseren führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen.

[0073] Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellkultursystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

Beispiel 3: Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

[0074] Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektiert, wobei hier jedoch selektionsbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [³H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunpräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d.h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

Beispiel 4: Herstellung und Anwendung eines HCV-RNA Konstrukt, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist.

[0075] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das einklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z.B. mittels Transfektion eines Ribozymkonstrucks oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende

Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

5

Beispiel 5: Kotransfektion eines HCV-RNA-Konstrukt mit einem separaten Luziferase-Transfektionskonstrukt

[0076] Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 (A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen ver-
10 mittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z.B. NS3-Protease-) Spaltungsstelle kodiert, mit einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen, insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen. Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der
15 Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt, die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestim-
20 mung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann ebenso gut ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmark-
ers ein Maß für die Replikation der viralen DNA.

25

Beispiel 6: Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als Leberzellspezifische Genfären für die Gentherapie

[0077] Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementie-
30 rende Helferzelllinien transfiziert, d.h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispiels-
weise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen
sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die
Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also
virus-ähnliche Partikel, die ein erfundungsgemäßes HCV-Subgenom-Konstrukt einschließlich des einklonierten Fremd-
35 gens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Kon-
strukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfundungsgemäßes HCV-
Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vor-
genannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine trans-
komplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als
40 Expressionsvektor.

Beispiel 7: Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte

(A) Isolationsverfahren

45

[0078] Für die Bestimmung adaptiver Mutationen und die Herstellung zellkultur-adaptierter HCV-RNA-Konstrukte wurde wie folgt verfahren: Zellen wurden mit einem HCV-RNA-Konstrukt wie unter den Beispielen 1 und 2 beschrieben transfiziert und G418-resistente Zellklone hergestellt. Zur Bestimmung der Replikationskompetenz (darunter wird in diesem Zusammenhang die Anzahl G418-resistenter Zellklone verstanden, die pro Mikrogramm transfizierter HCV-
50 RNA bzw. HCV-RNA-Konstrukt erhalten wird) wurde exemplarisch die Gesamt-RNA aus einem der Zellklone, genannt 9-13 (Fig. 1B, Spur 11) isoliert und die Menge der darin enthaltenen HCV-RNA mittels Northern-blot wie in Fig. 2 B beschrieben bestimmt. 10 Mikrogramm der Gesamt-RNA, die ca. 10^9 Moleküle HCV-RNA enthielt, wurde anschließend per Elektroporation in naïve Huh-7 Zellen eingeschleust (Fig. 9). Parallel dazu wurden 10^9 in vitro Transkripte der analogen neo-HCV-RNA, die mit isolierter Gesamt-RNA aus naïven Huh-7 Zellen auf eine Gesamt-RNA-Menge von 10 µg
55 aufgefüllt worden war, in naïve Huh-7 Zellen transfiziert. Nach Selektion mit G418 wurde die Anzahl der Zellkolonien, ausgedrückt in 'colony forming units (cfu) pro Mikrogramm RNA' in den beiden Ansätzen bestimmt. Bei einer Konzentration von 500 µg/ml G418 im Selektionsmedium betrug die Zahl der Kolonien, die mit der in der *isolierten* Gesamt-RNA enthaltenen HCV-RNA erhalten wurde, ca. 100.000 cfu pro Mikrogramm HCV-RNA. Dagegen wurden mit dersel-

ben Menge in vitro transkribierter HCV-RNA nur 30 - 50 Kolonien erhalten. Dieses Ergebnis belegt, daß die spezifische Infektiosität der HCV-RNA, die aus den Zellklonen isoliert wurde, ca. 1.000 - 10.000-fach höher ist als die Infektiosität der analogen in vitro Transkripte. Das methodische Vorgehen ist in Fig. 9 dargestellt.

- [0079] Mit Hilfe der 'long-distance RT-PCR' wurde die HCV-RNA aus der Gesamt-RNA der 9-13 Zellen amplifiziert, die PCR-Amplifikate kloniert und zahlreiche Klone sequenziert. Ein Vergleich der Sequenzen dieser reklierten RNAs mit der Sequenz der RNA, die ursprünglich in die naïiven Huh-7 Zellen eingeschleust wurde ergab, daß die reklierten RNAs zahlreiche Aminosäureaustausche besaßen, die über die gesamte HCV-Sequenz verteilt waren (Fig. 10). *Sfil*-Fragmente dieser reklierten Mutanten wurden im Austausch gegen das analoge *Sfil*-Fragment des ursprünglichen Replikikonstrukts in dieses eingeführt und RNAs der jeweiligen Mutanten wurden in naïve Huh-7 Zellen eingeschleust. Nach Selektion mit G418 wurde dann für jede HCV-RNA-Mutante die Zahl der gebildeten Kolonien bestimmt. Während mit der Ausgangs-RNA nur 30 - 50 Kolonien pro Mikrogramm RNA erhalten wurde war die Koloniezahl bei zwei der reklierten Varianten deutlich höher (Fig. 10). Im Fall der HCV-RNA-Konstrukte 9-13I und 9-13C betrug die spezifische Infektiosität 100 - 1.000 cfu pro Mikrogramm RNA und beim 9-13F Replikon sogar 1.000 - 10.000 cfu pro Mikrogramm RNA. Diese Ergebnisse zeigen, daß die Aminosäureaustausche in dem analysierten NS3-5B-Bereich der Mutanten 9-13I, 9-13C und insbesondere 9-13F zu einer deutlichen Erhöhung der Replikationskompetenz führten. Demgegenüber waren alle anderen HCV-RNA-Konstrukte (9-13 A, B, G, H und K) nicht mehr replikationskompetent, enthielten also letale Mutationen.
- [0080] Zwecks Beantwortung der Frage, welche der Aminosäureaustausche im 9-13F-Konstrukt zur Steigerung der Replikation führten, wurden die Austausche einzeln oder in Kombination in das Ausgangs-HCV-RNA-Konstrukt eingeführt und die entsprechenden RNAs in naïve Huh-7 Zellen eingeschleust. Das Ergebnis der Transfektionen mit diesen RNAs ist in Tabelle 1 zusammengefaßt. Daraus wird ersichtlich, daß im vorliegenden Beispiel die hohe Replikationskompetenz durch mehrere Mutationen bedingt ist. Den größten Beitrag leisten die Aminosäureaustausche in den HCV-RNA-Abschnitten NS5A und NS4B. Auch die einzelnen Austausche in der NS3-Region leisten einen Beitrag, der möglicherweise auf einem Synergismus dieser Einzelaustausche beruht.
- Diese Befunde belegen, daß es durch die G418-Selektion der Zellen, die mit den *neo*-HCV-RNA-Konstrukten transfiziert wurden, zur Anreicherung solcher HCV-RNAs kam, die eine deutlich höhere Replikationskompetenz hatten. Mit dem hier beschriebenen Versuchsansatz lassen sich HCV-RNA-Konstrukte mit sehr unterschiedlicher Replikationseffizienz selektieren. Je höher die Konzentration des Antibiotikums in dem Selektionsmedium ist, in/auf dem die HCV-RNA-Konstrukt-haltigen Zellen zwecks Selektion kultiviert werden, desto höher muß der Grad an adaptiven Mutationen und damit die Replikationseffizienz in den betreffenden HCV-RNA-Konstrukten sein, damit die Zellen auswachsen können. Werden die Selektionen mit niedrigeren Antibiotikum-Konzentrationen durchgeführt, können auch solche Zellen überleben und sich vermehren, die im Vergleich geringer adaptive Mutationen und eine weniger hohe Replikationseffizienz aufweisen.
- Das bisher beschriebene HCV-RNA-Konstrukt 9-13F, das mehrere adaptive Mutationen enthielt, hatte eine erwiesenermaßen höhere Replikationseffizienz als die parentale HCV-RNA. Um HCV-RNAs mit noch höherer Replikation in Zellkultur zu erhalten, wurde die HCV-RNA, die in der Gesamt-RNA eines ausgewählten Zellklons enthalten war, mehrfach in naïven Huh-7 Zellen passagiert. Dieser ausgewählte Zellklon, genannt 5-15, wurde durch Transfektion mit dem HCV-RNA-Konstrukt I₃₈₉/NS3-3' erhalten (Fig. 1). Er entspricht weitgehend dem Zellklon 9-13, der durch Transfektion mit einem HCV-RNA-Konstrukt hergestellt wurde, das eine um 22 Nukleotide kürzere HCV-IRES besaß (I₃₇₇/NS3-3'; Fig. 1). 10 Mikrogramm Gesamt-RNA, isoliert aus dem Zellklon 5-15, wurden mittels Elektroporation in naïve Huh-7 Zellen eingeschleust und die Zellen einer Selektion mit 1 mg/ml G418 unterzogen. Aus einem der so erzeugten Zellklone wurde wiederum Gesamt-RNA isoliert, in naïve Huh-7 Zellen transfiziert und analog selektiert. Dieser Vorgang wurde insgesamt viermal wiederholt. Nach der vierten Passage wurde aus einem Zellklon die Gesamt-RNA isoliert und die *neo*-HCV-RNA mit Hilfe der 'long-distance RT-PCR' amplifiziert. Das amplifizierte DNA-Fragment wurde mit dem Restriktionsenzym *Sfil* verdaut und in das *Sfil*-restringierte Ausgangskonstrukt I₃₈₉/NS3-3' inseriert. Insgesamt wurden über 100 DNA-Klone erhalten und zunächst mittels Restriktionsverdau analysiert. In vitro transkribierte RNA von ca. 80 dieser Klone wurde jeweils in naïve Huh-7 eingeschleust und einer Selektion mit 500mg/ml G418 unterzogen. Von den 80 untersuchten *neo*-HCV-RNA-Varianten erwiesen sich die allermeisten als replikationsdefekt. Bei zwei Mutanten, genannt 5.1 und 19, war die spezifische Infektiosität, ausgedrückt als 'colony forming units' pro Mikrogramm RNA, jedoch sehr deutlich erhöht (Tabelle 2). Durch mehrfache Passage der RNA in Zellkultur lassen sich offensichtlich HCV-RNAs herstellen, deren Replikationseffizienz aufgrund von Mutationen (sog. "adaptiven Mutationen") mehrere Größenordnungen höher ist als die der ursprünglich aus dem Patienten klonierten RNA.

(B) Modifikationsverfahren

- [0081] Solche nach (A) erzeugten und identifizierten adaptiven Mutationen können in ein wenig replikationskompetentes HCV-RNA-Konstrukt übertragen werden und führen zu einer massiven Steigerung der Replikation dieses Konstrukts. Diese Steigerung ist so hoch, daß damit nachweislich HCV-RNAs in Zellkultur zur Replikation gebracht

werden können, die kein selektierbares Markergen mehr besitzen. Fig. 12 zeigt einen Vergleich der Replikationseffizienz von HCV-RNAs, die entweder der Ausgangssequenz oder den adaptierten Sequenzen 9-13F bzw. 5.1 entsprachen. Zwecks einfacher Messung wurde das neo-Gen entfernt und durch das Gen für die Luziferase ersetzt. Als Negativkontrolle diente wiederum ein HCV-RNA-Konstrukt, das auf Grund einer inaktivierenden Mutation der NS5B RNA-Polymerase replikationsdefekt war. Schon 24 Stunden nach der Transfektion erkennt man einen deutlichen Unterschied in der Luziferaseaktivität zwischen der defekten RNA und den 9-13F bzw. 5.1-Konstrukten während zwischen der defekten RNA (318 DN) und dem Ausgangs-RNA-Konstrukt (wt) das keine adaptiven Mutationen besaß, kaum ein Unterschied zu sehen war. Während des gesamten Beobachtungszeitraums wurde die höchste Luziferaseaktivität und damit die höchste Replikation mit der 5.1-RNA erhalten. Diese Befunde belegen nicht nur die hohe Replikationseffizienz dieser RNA, sondern zeigen auch, daß es möglich ist, mit adaptierten HCV-RNA-Konstrukten ein Zellkultursystem aufzubauen, für das die Anwesenheit eines selektierbaren Gens nicht mehr notwendig ist. Eine zusammenfassende Übersicht der Nukleotid- und Aminosäureunterschiede zwischen dem Ausgangskonstrukt und den Mutanten 9-13F, 5.1 und 19 ist in Tabelle 3 gegeben.

15 Beispiel 8: Herstellung zellkultur-adaptierter HCV-RNA-Vollängengenome

[0082] In den Beispielen 1 bis 7 wurde stets eine subgenomische HCV-RNA verwendet, die die gesamte Strukturproteinregion von Core bis einschließlich p7 bzw. NS2 fehlte. Im vorliegenden Beispiel 8 wird gezeigt, daß es möglich ist, mit Hilfe der adaptierten NS3-5B-Sequenz ein HCV-Vollängengenom in Zellkultur zur Replikation zu bringen. Zu diesem Zweck wurde zunächst das SfiI-Fragment der gemäß Beispiel 7 hergestellten, hoch adaptierten HCV-RNA 5.1 in ein selektionierbares HCV-Vollängengenom transferiert (Fig. 12). Dieses HCV-Genom wurde in naïve Huh-7 Zellen transfiziert und einer Selektion mit unterschiedlichen G418-Konzentrationen unterzogen. In Abhängigkeit von der Selektionsstärke (der G418-Konzentration) wurde eine unterschiedlich große Zahl an Zellklonen erhalten (Fig. 12 B). Im Vergleich dazu wurden mit dem unveränderten HCV-Vollängengenom, das keine-adaptiven Mutationen enthielt, keine Kolonien erhalten, ebenso mit der Negativkontrolle, die auf Grund einer inaktivierenden Mutation in der NS5B RNA-Polymerase replikationsdefekt war. Zum Nachweis dafür, daß die so entstandenen Zellklone tatsächlich ein autonom replizierendes HCV-Vollängenkonstrukt enthielten, wurde Gesamt-RNA aus mehreren Zellklonen isoliert und mittels Northern-Blot analysiert. In allen Zellklonen war die Vollängen HCV-RNA eindeutig nachweisbar (Fig. 12). Damit ist eindeutig belegt, daß es mit Hilfe der an Zellkulturen adaptierter HCV-Sequenzen möglich ist, ein HCV-Vollängengenom herzustellen, das mit hoher Effizienz und autonom in einer Zelllinie repliziert, d.h. es können mit dem erfundungsgemäß System auch adaptierte HCV-Vollängengenome hergestellt werden. Da dieser Klon darüber hinaus die vollständige HCV-Sequenz besitzt, also auch die für die Viruspartikelbildung notwendigen Strukturproteine, ist es mit diesem System möglich, große Mengen infektiöser Viruspartikel in Zellkulturen herzustellen. Zum Nachweis dieser Viren werden zellfreie Überstände von Zellen, die ein replizierendes HCV-Vollängengenom tragen, auf naive Huh-7 Zellen gegeben und die so infizierten Zellen einer Selektion mit G418 unterzogen. Jeder Zellklon, der unter diesen Bedingungen auswächst, geht auf eine infizierte Zelle zurück. Die Viren in den Zellkulturüberständen von Zellen, die ein replizierendes HCV-Vollängengenom besitzen, können aber auch mit verschiedenen im Stand der Technik bekannten Verfahren wie Ultrazentrifugation oder Mikrodialyse angereichert und gereinigt werden und dann zur Infektion naiver Zellen verwendet werden. Mit diesem Verfahren ist eindeutig gezeigt, daß mit dem erfundungsgemäß HCV-Zellkultursystem zellkultur-adaptierte HCV-Vollängengenome hergestellt werden können, die mit hoher Effizienz in Zellen replizieren und infektiöse Viren produzieren. Diese können ebenfalls durch Infektion eines Versuchstiers, vorzugsweise dem Schimpansen, nachgewiesen werden.

Beispiel 9: Herstellung von HCV-Vollängen-Konstrukten und HCV-Subgenom-Konstrukten mit Reportergen.

[0083] Es wird ein HCV-RNA-Konstrukt hergestellt, bei dem anstelle des Antibiotikumresistenzgens ein Reportergen eingefügt wird (Fig. 13). Dabei kann die Replikation anhand der Menge bzw. der Aktivität des Reportergens bzw. Reportergenprodukts bestimmt werden. Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green fluorescence protein Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (sezernierte alkalische Phosphatasegen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z.B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z.B. enzyme-linked immunosorbent assay, ELISA) bestimmt werden. Das Reportergen kann entweder von einer eigenen IRES exprimiert werden oder in Form eines Fusionsproteins, das entweder als solches aktiv ist oder mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV-)Protease von diesem abgespalten wird.

Beispiel 10: Herstellung von HCV-Vollängen-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfahre für die Gentherapie oder als Expressionsvektor.

[0084] Das Konstrukt (Fig. 14) wird in Zellen eingeschleust und führt dort zur Bildung von HCV-Viruspartikeln, die zur Infektion weiterer Zellen verwendet werden können. Da die Viruspartikel eine RNA mit einem Fremdgen enkapsuliert haben kann dieses in den so infizierten Zellen zur Produktion des von diesem Fremdgen kodierten Proteins benutzt werden. Zellen, die mit dem Konstrukt transfiziert wurden, exprimieren ebenfalls das Fremdgen.

Beispiel 11: Herstellung von monocistronischen HCV-RNA-Konstrukten, bei denen das Resistenzgenprodukt als Fusionsprotein mit dem HCV-Anteil exprimiert wird.

[0085] Für bestimmte Untersuchungen ist es von Vorteil, wenn das HCV-RNA-Konstrukt kein heterologes IRES-Element besitzt. Solche Untersuchungen sind beispielsweise die Bestimmung der Interferonresistenz. Wird eine Zelle, die ein HCV-RNA-Konstrukt besitzt, mit Interferon-alpha oder -beta inkubiert, kommt es zu einer Reduktion der Replikation der HCV-RNA. Zur Aufklärung des Wirkungsmechanismus ist es notwendig, daß das HCV-RNA-Konstrukt keine heterologe IRES besitzt, da ansonsten nicht bestimmt werden kann, ob die Interferon-vermittelte Hemmung durch eine Hemmung der HCV-Replikation oder durch eine Hemmung der heterologen IRES vermittelt wird. Deshalb werden Konstrukte hergestellt, bei denen das Resistenzgen mit einem HCV-Protein fusioniert wird (Fig. 15). Entweder das Fusionsprotein ist als solches aktiv oder das Resistenzgenprodukt wird mittels einer proteolytisch spaltbaren Aminosäuresequenz so mit einem HCV-Protein verbunden ist, daß es von einer zellulären oder viralen (HCV)-Protease von diesem abgespalten wird.

Tabelle 1

Aminosäureaustausch ¹	HCV-Protein	cfu/ μ g RNA ²
kein		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	500 - 1000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5B	30 - 60
alle zusammen		5000

¹ Aminosäureaustausch im Polyprotein des HCV-Isolats Con-1 (EMBL-Genbank No. AJ238799); die Aminosäuren sind im Dreibuchstabenkode angegeben.

² Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

Tabelle 2

Spezifische Infektiositäten (cfu/ μ g RNA) des parentalen HCV-RNA-Konstruktls l ₃₈₉ /NS3-3'/wt und der Varianten 9-13C, 9-13I, 9-13F, 5.1 und 19.		
Transfizierte RNA-Variante	cfu/ μ g RNA ¹	
Wildtyp	30 - 50	
9-13 C	100 - 1.000	
9-13 I	100 - 1.000	
9-13 F	1.000 - 10.000	
5.1	50.000 - 100.000	
19	50.000 - 100.000	

¹ Colony forming units (Anzahl der Zellklone) bei einer Selektion von 500 μ g/ml G418.

5

10

15

20

25

30

35

40

45

50

55

Tabelle 3: Nukleotid- und Aminosäureunterschiede zwischen dem parentalen HCV-
RNA-Konstrukt I₃₈₉/NS3-3'/wt und den Mutanten 9-13I, 9-13F, 5.1 und 19

HCV Mutante	Nukleotidposition	Nukleotidaustausch	Aminosäureaustausch
9-13 I	3685	C > T	Pro > Leu
	4933	C > T	Thr > Met
	5249	T > C	-
	8486	C > T	-
	8821	G > A	Trp > stop
	8991	C > G	Arg > Gly
	9203	A > G	-
	9313	T > C	Phe > Ser
	9346	T > C	Val > Ala
	3866	C > T	-
9-13 F	4188	A > G	Arg > Gly
	4489	A > C	Glu > Ala
	4562	G > A	-
	4983	T > C	-
	5071	A > G	Lys > Arg
	5166	A > G	Lys > Glu
	6147	C > T	Pro > Ser
	6829	A > G	Glu > Gly
	7329	A > G	Lys > Glu
	7664	A > G	Ile > Val
NK5.1	8486	C > T	-
	8991	C > G	Arg > Gly
NK5.1	4180	C > T	Thr > Ile
	4679	C > T	-

5	4682	T > C	-
10	5610	C > A	Leu > Ile
15	6437	A > G	-
20	6666	A > G	Asn > Asp
25	6842	C > T	-
30	6926	C > T	-
35	6930	T > C	Ser > Pro
40	7320	C > T	Pro > Ser
45	7389	A > G	Lys > Glu
50	NK19	A > G	Glu > Gly
55	3946	C > G	Ala > Gly
60	4078	C > T	Thr > Ile
65	4180	T > C	-
70	4682	C > A	Leu > Ile
75	5610	A > T	Met > Leu
80	5958	T > A	-
85	6170	G > A	-
90	6596	C > G	Ala > Gly
95	6598	C > T	-
100	6833	C > T	-
105	6842	T > C	-
110	6930	A > G	Ser > Pro
115	7141	C > T	Glu > Gly
120	7320	A > G	Pro > Ser
125	7389	G > A	Lys > Glu
130	7735	G > A	Ser > Asn

[0086] Angegeben sind die Unterschiede der Nukleotid- und Aminosäuresequenzen zwischen der Ausgangs-HCV-
RNA-Sequenz Con 1 (EMBL-Genbank No. AJ238799) und denen der zellkulturadaptierten HCV-RNAs. Die Zahlen
beziehen sich auf die Nukleotid- und Aminosäurepositionen des HCV-Isolats Con1.

SEQUENCE LISTING

5 <110> Bartenschlager, Ralf
 <120> Hepatitis C Virus cell culture system
 <130> ba-1
 10 <140> 199 15 178.4
 <141> 1999-04-03
 <160> 11
 <170> PatentIn Ver. 2.1
 15 <210> 1
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 20 <400> 1
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cttccaggac 120
 ccccccctccc gggagagcga tagtggctcg cggaaaccgt gatcaccgg 180
 gacgaccggg tccttcttg gatcaaccgg ctcaatgcct ggatgtttgg gctgtcccc 240
 gcgagactgc tagccgatgt gtgttgggtc gggaaaggcc ttgtgttact gctgtatagg 300
 gtgttgcga gtccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 ctaaaagaaa aaccaaacct aacaccAACG ggcgcggcat gattgaacaa gatggattgc 420
 acgeaggttc tccggccgtc tggtggaga ggctattcgg ctatgactgg gcacaacaga 480
 caatcgctg ctctgatgcc gccgttgcgc ggctgtcagc gcaaggggcgc ccggttctt 540
 ttgtcaagac cgacctgtcc ggtgcctgtc atgaactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacggc gtccctgtcg cagctgtgtc cgacgttgc actgaagcgg 660
 gaaggactg gctgtatgg ggcgaagtgc cggggcaggta tccctgtca ttcacctt 720
 ctctgccga gaaagtatcc atcatggctg atgaatgcg gggctgtcat acgcttgatc 780
 cggtctacgt cccatcgac caccaagcga aacatcgcat cgacgagca cgtactcgga 840
 tggaaagccgg tcttgcgtat caggatgatc tggacgaaga gcatcagggg ctcgcggcag 900
 ccgaactgtt cgccaggctc aaggcgcgc tggccgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgc aatacatgg tggaaaatgg cccgttttgc ggattcatcg 1020
 actgtggccg gctggatgtc gggacccgt atcaggacat gggcttgcg acccgtgata 1080
 ttgtcagaagatgttgcgc gaaatggctg accgttctt cgtgttttac ggtatcgccg 1140
 ctcccgattt ctagcgatc gcttctatc gcttcttgc cgagtttttc tgagttaaa 1200
 cagaccacaa cggtttccct ctageggatg caatccgcgc ccttcctccccc cccccccct 1260
 35 aacgttactg gccgaagccg tggaaataaa gggccgtgtc gttttgtata tatgttattt 1320
 tccacccat tggccgttgc tggcaatgtg agggccggaa aacatggccc tggccgttgc 1380
 acgagcatc ctagggtct tccctctc gccaaggaa tgcaaggatc gttgaatgtc 1440
 gtgaagggaaatgttgc ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcgc ggaacccccc acctggcgac aggtgcctct gggccaaaaa gccacgtgt 1560
 taagatacac ctgcaaaaggc ggcacaaaccc ctaggtccacg ttgtgagttg gatagtttg 1620
 gaaaggtca aatggcttc ctaaagcgta ttcaaaagg ggtcggagga tgcccaag 1680
 gtacccatt gtatggatc tgatctgggg cctcggtcga catgtttac atgttttag 1740
 tcgaggtaaa aaaacgtcta ggcggccgc accacggggc cgtggtttgc ctttgaaaaa 1800
 cacgataata ccatgggcac gaatcttaaa cctccaaagaa aaacccaaacg taacacccaa 1860
 cgccggccac aggacgttca gttccggc ggtgttgcga tggccgttgc agtttacctg 1920
 ttggccgcga gggcccccag gttgggtgtc cggccgacta ggaagacttc cgagcggtcg 1980
 40 caacctcgta gaaggcgaca acctatcccc aaggctcgcc agcccgagg tagggcctgg 2040
 gtcggcccg ggtacccctg gccccttat ggcataatggg gttgggggtc ggcaggatgg 2100
 ctctgtcact cccgtggcgc tggccatgt tggggccca cggacccccc gctgtggcgc 2160
 cgcaatttgg gtaaggatcat cgatccctc acgtgcggct tggccgtatc catgggttac 2220
 attccgtctc tggccgcctt cctagggggc gtcggccaggcc ccttggcgc tggccgtccgg 2280
 gtttggagg acggcgtaa ctatgcacca gggatctgc cgggttgc tttttctatc 2340
 ttcttttgg ctgttgcgtc ctgtttggacc atcccgatgtt cccgttatga agtgcgcac 2400
 gtatccggag gttaccatgtt cacaacgc tgcctcaacg caagcattgtt gatgaggca 2460
 45 ggggacatgtt ctagcatac cccgggtgc tgcctgtcg ttcggggagaa caactccccc 2520
 cgctgtggg tagcgtctac tccacgtc gggccaggaa acgttagcgt ccccaactacg 2580
 acgataacgac gccatgtcga ttgtctcgat gggccggctg ctctgtcgc cgctatgtac 2640
 gtgggagatc tctgcggatc tttttctc gtcggccagc tggccacatt ctgcgtcc 2700

55

	acgctgcacg	ggccaaacgc	cctgtgtat	aggctggag	cggtcaaaa	cgaggtaact	6720
5	accacacacc	ccataaccaa	atacatcatg	gcatcatgt	cgctgaccc	ggaggtegtc	6780
	acgagcacct	gggtgttgtt	aggcggagtc	ctagcagctc	tgcgcgcgt	ttgcctgaca	6840
	acaggcagcg	tggtcatgtt	ggcaggatc	atcttgcgg	gaaagccggc	catcatccc	6900
	gacagggaa	tccttaccc	ggagttcgat	gagatggaa	agtgcgcctc	acacccccc	6960
	tacatcgaa	agggaaatgc	gctcgccaa	caatcaa	agaaggcaat	cgggttgtcg	7020
	caaacagcca	ccaagcaagc	ggaggctgt	gtccccgtt	tgaatccaa	gtggcggacc	7080
10	ctcgaagcct	tctggggcga	gcatatgtt	aatttcatc	gcccgtata	atatttagca	7140
	ggcttgtcca	ctctgcctgg	caaccccg	atagatcac	tgatggcatt	cacagcctct	7200
	ataccaggcc	cgctcaccac	ccaacatacc	ctcccttta	acatcctggg	ggatgggtg	7260
	gcccacca	ttgcgtctcc	cagcgtctgt	tctgtttcg	taggcgcgg	catcgctgg	7320
	gcccgttgg	ccagcatagg	cctttggaa	gtgttgtt	atatttgttgc	aggttatgg	7380
15	gcagggttgg	caggcgcgt	cgtggccctt	aaaggatcg	gcccgcagat	gcccctccacc	7440
	gaggacctgg	ttaacctact	ccctgtata	ctctccctgt	gcccctagt	cgtcggggtc	7500
	gtgtgcgcag	cgatactgc	tcggcaatgt	gcccagggg	agggggctgt	gcagtggatg	7560
	aaccgctga	tagcgttcgc	ttcgcgggtt	aaccacgtt	ccccacgc	ctatgtgcct	7620
	gagagcgc	ctgcgcac	tgtcaactcg	atccctcta	gtcttacat	cactcagctg	7680
20	ctgaagggc	ttcaccagg	gatcaacag	gactgttca	ccggatgtc	cggctgtgg	7740
	ctaagagatg	tttgggattt	gatatgcac	gtgttgtact	atttcaagac	ctggctccag	7800
	tccaagctcc	tgccgcgtt	gcccggatc	cccttcttct	catgtcaac	tgggtacaag	7860
	ggagtctggc	ggggcgcac	catcatcaa	accacgtcc	catgtggac	acagatcacc	7920
	ggacatgtga	aaaacggttt	catgaggat	gtggggcctt	gacacgttgg	taaacacgttgg	7980
	catggacat	tccccattaa	cgcgtacacc	acggggccct	gacgccttcc	cccgccgc	8040
25	aattattcta	ggggcgtgt	gccccgtgt	gtgttgttgc	acgtggaggt	tacgcgggt	8100
	ggggatttcc	actacgtac	gggcatgacc	actgacaatgc	taaagtgc	gtgtcaggtt	8160
	ccggcccccc	aatttttcac	agaagtggat	gggggtgggt	tgcacaggtt	cgctccagcg	8220
	tcaaaaaacc	tcctacggg	gggggttca	ttccctgttgc	ggctcaatca	atacctgggtt	8280
	gggtcacagc	tcccattgcg	gcccgaacc	gacgtacag	tgctcaatc	catgtcaacc	8340
	gaccctcc	acattacgg	ggagacggct	aagcttaggc	tggccagggg	atctcccccc	8400
30	tccctggcca	gctcatc	cgtccatgt	tctgcgtt	cccttgcggc	aacatgca	8460
	acccgtcatg	actccccc	cgctgaccc	atcgaggcc	actcttgc	cgccgcaggag	8520
	atgggggg	acatcaccc	cggtggatc	aaaaatagg	tagtatttt	ggactctttc	8580
	gagccgc	aagcggag	ggatgagagg	gaagtatcc	ttccggcg	gatctgtgg	8640
	agg	tttccctcg	agcgtatccc	atatgggcac	gcccggatta	caaccctcca	8700
35	ctgttagat	cctggaa	cccggtact	gtccctcc	tgttacac	gtgtccattt	8760
	ccgcctggca	aggccctcc	gataccac	ccacggagg	agaggacgtt	tgtctgtca	8820
	gaatctacc	tgtctctgc	tttggcgag	ctgcaccaa	agacccctgg	cagctccgaa	8880
	tcgtcgcc	tcgacag	cacggcaac	gccttcttgc	accagccctc	cgacgcacgg	8940
	gacggggat	ccgacgtt	gtcgtact	tccatgccc	cccttggagg	ggagccgggg	9000
	gatcccgat	tcgacgtac	gtcttgcgt	accgttgc	aggaggctgt	tgaggacgtc	9060
	gtctgtgt	cgatgttca	catatggca	ggccgttgc	tcacgcatt	cgctgcggag	9120
	gaaacca	tgccttcat	tgcactgac	aactttgc	tcctgtacca	caacttgg	9180
40	tatgtctaa	catctcg	cgcaaggct	cgccagaaga	agtcac	tgacagact	9240
	cagg	acgacca	ccgggac	ctcaaggaga	tgaaggc	ggcgtccac	9300
	gttaagg	aacttctat	cgtggagaa	gcctgtat	tgaccccc	acatttggcc	9360
	atgat	tttccat	ttggatgtt	ggcaaggac	tatccag	ggccgttaac	9420
	cacatcc	ccgtgt	ggactgt	gaagac	agacacaa	tgacaccacc	9480
	atcatgg	aaaatggat	tttctgc	caacc	agggggcc	caagccag	9540
45	cgccttat	tattcc	tttgggg	cgtgtgt	ggaaaatgg	cctttacat	9600
	gttgtctt	cccttcc	ggccgtat	ggcttctat	accgttcc	atacttctt	9660
	ggacagc	tcgat	gttgaat	tttgggg	tttggctt	tatgggctt	9720
	gcatatg	cccgctt	tttgc	gttgc	tttgggg	tgttggag	9780
	tcaatctacc	aatgtt	tttgg	tttgg	tttgggg	gtcgtcaca	9840
	gagcgtt	atcatgg	gggg	tttgg	tttgggg	cggttat	9900
	cgtgtcc	cgacgtt	tttgg	tttgg	tttgggg	atgttactt	9960
	aaggccgt	cggttgc	tttgg	tttgg	tttgggg	cgatgtcg	10020
50	gacgac	tcgttat	tttgg	tttgg	tttgggg	tttgggg	10080
	gccttac	aggctat	tttgg	tttgg	tttgggg	tttgggg	10140
	tacgtat	tttgg	tttgg	tttgg	tttgggg	tttgggg	10200
	ggcaaa	tgtat	tttgg	tttgg	tttgggg	tttgggg	10260
	gagac	tttgg	tttgg	tttgg	tttgggg	tttgggg	10320
	acc	tttgg	tttgg	tttgg	tttgggg	tttgggg	10380
	caact	tttgg	tttgg	tttgg	tttgggg	tttgggg	10440
	gac	tttgg	tttgg	tttgg	tttgggg	tttgggg	10500
	tctcc	tttgg	tttgg	tttgg	tttgggg	tttgggg	10560
	cgag	tttgg	tttgg	tttgg	tttgggg	tttgggg	10620

5 gctgccactt gtggcaagta cctttcaac tggcagtaa ggaccaagct caaactcact 10680
ccaatcccg ctgcgtcccc gttggattta tccagctgg tcgttgcggg ttacagcggg 10740
ggagacatat atcacagcct gtctcggtcc cgaccccgct gttcatgtg gtgcctactc 10800
ctacttctg tagggtagg catctatcta ctcccccaacc gatgaacggg gagctaaaca 10860
ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt tttttttttt 10920
ttttttttt ttttttttc tcctttttt ttccctttt ttccctttc ttccctttgg 10980
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
agagagtgtct gatactggcc tctctgcaga tcaagt 11076

10

15

20

25

30

35

40

45

50

55

<210> 2
 <211> 8637
 <212> DNA
 <213> Hepatitis C Virus

5 <400> 2
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 ttttacgc gaaagcgct agccatggc ttagtatgag tgcgtgcag cttccaggac 120
 cccccctcc gggagagcca taggtgtctg cgaaaccgg gagaatccac gattggccag 180
 gacgacccgg tccttttgc gatcaacccg ctcaatgcct ggagattttgg gctgtcccc 240
 10 gcgagactgc tagccgatg gtgttggtc gcgaaaggc ttgtgtact gcctgatagg 300
 gtgttcgcg gtggccccc aggctctgt gaccgtgcac catgagcag aatcctaaac 360
 ctcggaaa aaccaaaaggc cgcgcctatg tgaacaaga tggattgcac gcaggttctc 420
 cggccgttg ggtggagagg ctattccgc atgactggc acaacagaca atcgctgct 480
 ctgtatgcgc cgtgttccgg ctgtcagcgc agggcgcgg ggttctttt gtcagaccc 540
 acctgtccgg tgccctgatg gaactgcagg acgaggcage gggctatcg tggctggca 600
 15 cgacggccgt tccttcgc gctgtgc gctgtgtc acgttgcac taaaggcggg aggactggc 660
 tgctattttgg cgaagtgcgg gggcaggatc tcctgtatc tcaaccttgc cttccggaga 720
 aagtatccat catggctatg gcaatgcggc ggctgcatac gtttgcattc gtcacactgc 780
 cattcggcca ccaagcgaaa catcgatcg aegcagcag tactcgatg gaagccggc 840
 ttgtcgtatc ggtatgtctg gacgaagagc atcaggggc cgcgcggc gaaactgttcg 900
 ccaggctcaa ggcgcgc gcccggcagg aggttgcgt cgtgacccat ggcgatgc 960
 20 gcttgcggaa tatcatggt gaaatggc gttttctgg attcatgc gtcgtatc gttggccggc 1020
 tgggtgtggc ggaccgtatc caggacatag cggttgc tccgtatatt gtcgtatc 1080
 ttggcggcga atgggctgac cgcttcctcg tgcttacgg tatecgccgc cccgattcgc 1140
 agccatcgcc ctctatcg cttcttgcg agtttcttgc agttttaaca gaccacaacg 1200
 gttccctct acggggatca attcggccccc ttcctccccc cccccccta cgttactggc 1260
 cgaagecgct tggatataagg cgggtgtcgt tttgtatata tttatatttgc cccatattt 1320
 25 ccgtctttg gcaatgttag gggccggaa cctggccctg ttccttgc gggcattct 1380
 aggggtcttt cccctctcg caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
 gtctctctgg aagcttcttgc aagacaaaca acgtctgtcg cggccctttt cggcagccg 1500
 aaccccccac ctggcgacag gtgcgttcg ggcggaaaacg cactgtata agatacacct 1560
 gcaaaaggcg gcaaaacccca gtcacccgtt gtggatggg tagtgggaa aagagtcaa 1620
 tgctcttc tcaagcttatt caaacagggg ctgaaggatg cccagaagggt accccatttt 1680
 atggatctg atctggggc tcggtgacca tgcttacat gtttttagt gaggtaaaa 1740
 30 aacgttctagg ccccccgaac cacggggacg tggtttctt ttggaaaaaca cgataataacc 1800
 atggacccggg agatggcagc atcgtgcggg ggcgggggtt tgcgtgtt gatactctt 1860
 accttgcac cgcactataa gtcgttcctc gtcgtgttca tatgtggg tttatattt 1920
 atcaccaaggc cggaggcaca cttgcgtactg tggatcccc ccccaacgt tcggggggc 1980
 cgcgtatccg tcatcttc tcaatgcgcg atccacccag agctaatttt taccatcacc 2040
 aaaatcttcg tcgcataact cggtccact atgggtcttcc aggtgtgtat aaccaaagt 2100
 ccgtacttcg tgcgcgcatac cgggtcttcc cgtgtatgc tgcgtgtc gaaagggtct 2160
 35 ggggttcatt atgtccaaat ggtctctatg aatggggccg cactgtacccg tacgtaccc 2220
 tatgaccatc tcaccccaat gggggactgg gcccacgcgg gcttacgaga ctttgcgggt 2280
 gcagttgagc cggcgtctt ctgtatgc gagaccaagg ttatcacctg gggggcagac 2340
 accggccgt gtggggacat catcttggc ctgcgtct cccggccgcg ggggaggag 2400
 atacatcttc gaccggcaga cggcttggc gggcgggggtt ggcgtactt cgcgcctatt 2460
 acggctact cccaaacggc gcgaggcata cttgcgtca tcatctact cctcacaggc 2520
 cgggacacca accagggtca ggggggggtt caaggtgttcc cccacccggc acaatcttc 2580
 40 ctggcgcact ggcgtatgg cgtgtttgg actgtctatc atgggtccgg ctcaaagacc 2640
 ctggccggcc caaaggggccc aatcacccaa atgtacacca atgtggacca ggacccgtc 2700
 ggctggcaag cggcccccgg ggcgcgtcc ttgacacat gacccgtcc cagctcgac 2760
 cttaacttgg tcacggacca tggatgttgc tccctgggtc gccggggggg cgacacagg 2820
 gggagctac tctcccccgg gcccgtctcc tacttgaagg gtccttcggg cggccactt 2880
 ctctggccct cggggcgcacg tggggcata tttccggctg cctgtgtc cccgggggtt 2940
 45 gcaaggccg tggacttttgc acggcgtatg totatggaaa ccactatgcg gtcccccggc 3000
 ttacggaca actcgcccccc tccggccgtt cggcagacat tccagggtggc ccatctacac 3060
 gcccctactg gtacggccaa gggcactaaat gtcgtggctt cgtatgcgc ccaagggtat 3120
 aagggtcttgc ttcttgcac gtcgtgtcc gccccttag gtttggggc gttatgtct 3180
 aaggcacatg gtatcgacc taacatcaga accggggtaa ggaccatcac cacgggtgcc 3240
 cccatcacgt actccacca tggcaagttt cttggcgcac gttgtgtctc tggggggc 3300
 50 tatgacatca taatatgttgc tgagtggccac tcaactgact cgaccactat cttgggcac 3360
 ggcacagttcc tggaccaaggc ggagacggct ggagcgcgc tgcgtgtct cggccaccgt 3420
 acgcctccgg gatcggtcac cgtggccatc ccaacatcg aggaggtggc tctgtccac 3480
 actggagaaa tcccttttgc tggcaagcc atcccatcg agggatcaa gggggggagg 3540

5	cacccatctt tctggcattc caagaagaaa tgtgtatgac tcggcgccaa gctgtccggc 3600 ctcggaactca atgtctgtac atattaccgg ggccttgatg tatccgtcat accaaactagc 3660 ggagacgtca ttgtctgtac aacggacgtc ctaatgcgg gctttaccgg cgatttcgc 3720 tcagtgtatc actgtcaatac atgtgtcacc cagacatgcg acttcagecct ggaccggacc 3780 ttcacccatc agacgacgac cgtgcaccaa gacgcgggtg caccgtcgca gggcgaggc 3840 aggactggta ggggcaggat gggcatttac aggttgtga ctccaggaga acggccctcg 3900 ggcatgttcg attcctcggt tctgtgcag tgctatgacg cgggctgtgc ttggtagcag 3960 ctcacggccc cggagaccc agttaggtt cgggcttacc taaacacacc agggttgcc 4020 gtctgcggg accatcttgc gttctgggg agcgtcttta caggcttcac ccacatagac 4080 gccccatttc tgcctccagac taagcaggca ggagacaact tcccctaccc tggatcatc 4140 caggctacgg tgcgcgcag ggctcaggct ccacccatc cgtgggacca aatgtggaa 4200 tgtctcatac ggctaaagcc tacgctgcac gggccaacgc ccctgtctga taggctggg 4260 ggcgctcaaa acgagggttc taccacaccc cccataacca aatacatcat ggcatgtac 4320 tcggctgacc tggagggtcg cagcagcacc tgggtgtcg tagggcggat cctagcagct 4380 ctggccgcgt attgcgtgcac aacaggcagc gtgttcattt gggaggatc catctgtcc 4440 gaaaaaggccg cccatcatcc cgacaggggaa gtcccttacc gggagttcgta tgagatggaa 4500 gagtgcgcct cacacccccc ttacatcgaa cagggaaatgc agtcgcgcg acaattccaa 4560 cagaaggca tcgggttgc gcaaaacagcc accaaagcaag cggaggctgc tgctccctcg 4620 gtggaaatcca agtggcgac ctcgaacgg ttctggcgta acatatgtg gaatttcatc 4680 agcgggatata aatatttgc aggttcgtcc acttcgtccgt ccaaccccgcc gatagcatc 4740 ctgatggcat tcacagcctc tatcaccacgc ccgtccacca ccaacacatac cctctgttt 4800 aacatccctgg ggggatgggt ggccggccaa ctgtcttc cagcgcgtc ttctgtttc 4860 gttagggccg gcatcgctt agcgggttgc ggacagcatag gcttgggaa ggtgttgc 4920 gatattttgc cagggttatgg agcagggtgt ggacgcgcgc tctgtgcctt taaggatcg 4980 agcgggagaa tgccctccac cgaggacatcg gttaaacttac tccctgtat cctccctcc 5040 ggcccttag tcgtcggtt cgtgtgcga gcgataactgc gtccgcacgt gggcccgagg 5100 gagggggctg tgcagtggat gaaccggctg atagcgttc cttcgccggg taaccacgtc 5160 tccccccacgc actatgtcc tgagacgcac gtcgcacac gtcgtactca gatctcttc 5220 agtcttacca tcaactcgat gctgaagagg ctttaccatg gatcaacga ggatgtcc 5280 acggccatgt ccggctcgat gctaagagat gtttggatt ggtatgcac ggtgttgc 5340 gatttcaaga cctggcttca gtccaaagtc ctggcgcat tgccgggagt ccccttctc 5400 tcatgtcaac gtgggtacaa gggagttcg cggggcgacg gcatcatgca aaccacctgc 5460 ccatgtggag cacagatcac cggacatgtg aaaaacgtt ccatggatc cttcgccgg 5520 aggacatgtg gtaacacgtg gcatgaaaca ttccccattt acgcgtacac cacggggcc 5580 tgcacggccc ccccggeggc aaattatttc agggtcgat gggcggttgc tgctggagg 5640 tacgtggagg ttacgcgggt gggggatttc cactacgtga cgggcgtac cactgacaac 5700 gtaaaagtgcg cgtgtcgat tccggcccccc gaatttctca cagaagtggta tgggtgcgg 5760 ttgcacaggat acgtccacg gtgcacaaaccctt ctccatcggg agggatgtac attctgtc 5820 gggtctcaatc aatacttgcg tgggtcagac cttccatcg cggccaaacc ggcgtactca 5880 gtgtctactt ccatgtccac cggaccctcc cacaattacgg cggagacggc taagctgtt 5940 ctggccaggg gatctccccctt ctccctggcc agtcgtatcg ctggcactg gtctgcgcct 6000 tccttgaagg caacatgcac taccgtcat gactccccgg acgtgtaccc catcgaggcc 6060 aaccctctgt ggccggcaggaa gatggggggg aacatcaccc gctgtggatgc agaaaataag 6120 gttagtaattt tggactttt cggccgtcc caagcgagg aggatgagaa ggaagtatcc 6180 gttccggcgag agatctcgat gagggtccagg aaattccctc gaggcatgc catatggca 6240 cgcccggtt acacccctcc actgttagag tccctggaaagg accccggacta cgtccctcca 6300 gtggtacacg ggtgtccatt gccgcctgc aaggccccctc cgataccacc tccacggagg 6360 aagaggacgc ttgtctgtc agaatcttac gtgtcttctg ctttggccgg gtcgtccaca 6420 aagacccatcg cggacttcac atgtcgcc gtcgacacgc gcacggccac ggccctcttc 6480 gaccggccct cggacgcacgg cgacgcggga tccggatgtt agtcgtactc ctccatgccc 6540 cccccttgagg gggagccggg ggatcccgat ctcaacgcac ggttctggcc taccgtaaagc 6600 gaggaggctt gtgaggacgt cgtctgtcc tctgtgttcc acacatggac aggegcctcg 6660 atcacgcctt cgcgtccggaa ggaaaccaac ctgcacccatca atgcactgtac caactcttc 6720 ctccgttacc acacacttgc tctatgtaca acatctcgca ggcacggccct gccgcggcaga 6780 aaggtaactt cggacacatc gagggtcgat gacggacactt accggggactt gtcacaaagg 6840 atgaaggcga aggcttccac agttaaggct aaacttctat ccgtggaggaa agcctgtaa 6900 ctgacggcccc cacattccgc cagatctaa tttggctatg gggcaaaaggaa cgtccggaa 6960 ctatcccgat cggccgttac ccacatccgc tccgtgtgg aggacttgc ggaagact 7020 gagacaccaa ttgacaccatc catcatggca aaaaatggg ttttgcgtt ccaacccagg 7080 aaggggggcc gcaaggccacg tccgcatttac gtttccccc atttggggat tctgtgtgc 7140 gagaaaatgg cccttacga tgggtctcc acctccctc agggcgat gggcttctca 7200 tacggattcc aatactctcc tggacacgcgg gtcgatgtcc tggtaatgc ctggaaagcg 7260 aagaatgtcc ctatgggtt cgcacatgcg accccgttgc ttgactcaac ggtactgt 7320 atgacatcc gttgtggatc gtcatactac caatgtgtg acttggcccc cgaaggccaga 7380 caggccatccat ggtgcgtccac agagcggtt cttatcgcccc gccccctgcac taatttca 7440 ggggcagaact gggcgatcccgcc cgggtgcgc gggcggtt tactgacgac cagctgtgggt 7500
10	
15	
20	
25	
30	
35	
40	
45	
50	

aataaccctca catgttactt gaaggccgct gcggcctgtc gagctgcga 7560
 tgcacgatgc tcgttatgcgg agacgacctt gtcgttatct gtgaaagcgc ggggacccaa 7620
 gaggacgagg cgagcctacg ggccctcacg gaggctatga ctagatactc tgccccccct 7680
 5 ggggaccgc ccaaaccaga atacgacttg gagttataa catcatgtc ctccaatgtg 7740
 tcagtcgcgc acgatgcatt tggcaaaagg gtgtactatac tcacccgtga ccccacccacc 7800
 ccccttgcgc gggctgcgtg ggagacagct agacacactc cagtcaattc ctggctaggc 7860
 aacatcatca tgtatgcgc caccttgg gcaaggatga tcctgtatgac tcatttcttc 7920
 tccatccttc tagtcagga acaacttga aaagccctag attgtcagat ctacggggcc 7980
 10 tggtaacttca ttgagccact tgacctaactc cagatcatc aacgacttca tggccttagc 8040
 gcattttcac tccatagttt ctctccaggt gagatcaata gggtgttcc atgcctcagg 8100
 aaacttgggg taccgcctt cgcgactctgg agacatcggtt ccagaagtgt ccgcgcctagg 8160
 ctactgtccc agggggggag ggctgccact tggcaagt accttttcaa ctgggcagta 8220
 15 aggaccaagc tcaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280
 ttcgttgcgt gttacagcg gggagacata tatcacagcc tgcgttgcgtc ccgacccgc 8340
 tggttcatgt ggtgcctact cctactttct gttagggtag gcatctatct actccccaac 8400
 cgatgaacgg ggagctaaac actccagcc aataggccat cctgtttttt tcccttttt 8460
 tttttttttt tttttttttt tttttttttt tttttttttt ctccctttttt tttccctcttt 8520
 ttttcctttt ctttcctttt gttggctccat cttagcccta gtcacggta gctgtgaaag 8580
 gtccgtgagc cgcttgactg cagagagtgc tgatactggc ctctctgcag atcaagt 8637

20

25

30

35

40

45

50

55

<210> 3
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus
 5
 <400> 3
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tttcacgca gaaagcgctc agccatggcg ttagtatgag tgctgtgcag cctccaggac 120
 ccccccctccc gggagagcca tagtggtctg cggaaacggg gatgacaccg gaattgcccag 180
 10 gacgaccggg tccttctt gatcaacccg ctcaatgcct ggagattttgg gcgtgcccc 240
 gcgagactgc tagccgagta gtgtgggtc gcgaaaaggcc ttgtgttaact gctgtatagg 300
 gtgtgtgcgta gtgcggccggg aggtctcgta gaccgtgcac catgacgacg aatccctaaac 360
 ctcaaaagaaa aaccaaacgt aacccaacg ggcgcgcct gattgaacaa gatggatttc 420
 acgcagggtc tccggccgct tgggtggaga ggcttattcg ctatgactgg gcacaacaga 480
 caatcggtg ctctgtatgcc gccgtgttcc ggctgtcgcgc 540
 ttgtcaagac cgacactgtcc ggtgcctgaa atgaaactgca ggacgaggca ggcgcgctat 600
 cgtggctgcgac cgcgcggg gttcctgtcg cagctgtctc catgacgacg aatccctaaac 660
 15 gaaggggactg gctgttattt ggcgaagtgc cggggcaggat ttcctgtca ttcacaccc 720
 ctctgtccgaa gaaagatattcc atcatgttgc atgaaatcgcc gggctgtcat acgcttgc 780
 cggttacccg cccattcgac caccacggca aacatcgcat cgagcgagca cgtactcgga 840
 tggaaaggccg tttgtcgat caggatgtatc tggacaaaga gcatcagggg ctcgcgc 900
 ccgaactgtt cggcggctc aaggcgcgcg tggccggacgg cgaggatctc gtctgtaccc 960
 20 atggcgatgc ctgttgcgg aataatcatgg tggaaaatgg cgcgttttgc ggattcatcg 1020
 actgtggccg gttgggtgtc gggggccgtt atcaggatgtc agcgttgcgtt acccggtata 1080
 ttgtcaaga gcttggccgc gaatgggtcg accgttctt cgtgttttac ggtatcgccg 1140
 ctcggatttgc acggcgcata gccttctatc gccttcttgc 1200
 cagaccacaa cggtttccct ctatcgccgtt caattccgcg cctctccctt cccccccctt 1260
 aacgttactg gccgaaggccg ctggaaataa ggcgggtgtg ctgtgttata tatgttattt 1320
 25 tccaccatgtt gggccgtt tggcaatgtg aaggccggaa aacgtggccg tgcgttcttgc 1380
 acgagcatcc ttaggggtctt ttcctctc gccaaaggaa tgcaagggtctt gttgaatgtc 1440
 gtgaaggaaag cagttccctt ggaagcttctt tgaagacaaa caacgtctgtt agcgcaccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgtccctt gccgcggaaa gccacgtgtt 1560
 taagatatacc ctgcggatcc ggcacaaccc cagtggccacg ttgtgagggtt gatagttgtt 1620
 gaaagagtc aatggcttc tcaacggta ttcacaaagg ggctgaaggaa tgcccaagaag 1680
 gtacccattt gtatggatc tgatctgggg ctcgggtgtca catgctttac atgtgttttag 1740
 30 tcgaggtaaa aaaacgtcta ggcggccggaa accacggggaa cgtggtttcc ttgtaaaaaa 1800
 cacgataaa ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
 tgcattatca ttagccctac aggccggac aggaacccagg tgcaggggaa ggtccaaatgtt 1920
 gtctccaccc caacacaaatc ttccctgggg acctgggtca atggcggtgtt ttggactgtc 1980
 ttcatgttgc cggcgtcaaa gacccttgcg gcccggaaagg gcccaatcac ccaaatagtac 2040
 accaatgtgg accaggaccc cgtcggttgg caagcgcccc cggggggccg ttccttgaca 2100
 35 ccatgcactt cggcgcgttcc ggaccctttac ggcatttgcg ggcatttgcg tgcatttccg 2160
 gtgcggccgc gggcgacagc caggggggacg ctactctcc ccaaggccgtt ctcctacttg 2220
 aagggtctt cggggcggttcc actgtcttcg ccctcggttcc acgctgttggg catctttccg 2280
 gctggcggttgc gaccccgagg ggttgcgaag ggcgggttgc ttgttacccgtt cgagtctatgt 2340
 gaaaccacta tgccgtcccc ggttccatcg gacaactcg tccctccggc cgtaccgcag 2400
 acatccagg tggcccatat acacggccctt actggtagcg gcaagagac taaagggtccg 2460
 gtcgtatgttgc gaccccaagg gtataagggtt ctgttgcgtt accccgttccgtt cggccggacc 2520
 40 cttaggtttcg gggcgatataatgttgcgcatgttgc acccttaatcatcagaacccggg 2580
 gtaaggacca tcaaccacggg tgcccccattt acgtacttca cctatggca gtttcttgc 2640
 gacgggtgtt gctctggggg cgcctatgtac atcataatat gtgtatggatg ccactcaact 2700
 gactcgacca ttcattccggg catcgccaca gtccttgcggc aaggcgagac ggctggagcg 2760
 cgactcggtcg tgcgtccatcg cgtacccgtt ccgggttgcg tcaccgttgc acatccaaac 2820
 atcgaggagg tggctctgtc cagcacttgcgaa gaaatccctt ttatggca agccatcccc 2880
 atcgagacca tcaagggggg gaggcacccatcatttgcgatcatttcaagaa gaaatgtgt 2940
 45 gagctcgccg cgaagctgtc cggccgttgcgatcatttgcgatcatttcaacttccatatttccggggcett 3000
 gatgtatccg ttcattccaaatc tagggagac gtcatttgcgatcatttcaacttccatatttccggggcett 3060
 acggggcttca cggcgattt cgtactcgatcgttgc acatgtgttccatccatatttccatatttccggggcett 3120
 gtcgacttca gcttgcggcc gacccttgcacc attggatcgatcgttgc acatccaaac 3180
 gtgtcacgttgc cgcacggccg aggccggactt ggttggggca ggttggccat ttcacccgtt 3240
 gtgactccatcg gagaacggcc ctcggccatcg ttcgttccctt cgttgcgttgc ggttggccat 3300
 50 gacggggctt gtcgttgcgatcgttgc cccggccgatcg ccttgcgttgc ggttggccat 3360
 tacctaaaca caccagggtt gccgttgc caggccatcg ttcgttgc ggttggccat 3420
 ttacccatcg ttcaccatcg agacccatcg ttcgttgc ggttggccat 3480
 aactccctt acctggatcg ataccatcg acgggttgcgatcgttgc cccggccatcg 3540

	ccatcggtgg accaaatgtg gaagtgtctc atacggctaa agcttacgct gcacgggca 3600
5	acgccccctgc tgtatacgct gggagccgtt caaaacagg ttactaccac acacccccata 3660
	accaaataca tcatggcatg catgtcgct gacccgggg tcgtcacgag cacctgggtg 3720
	ctggtaggcg gagtccttagc agctctggcc gctgtattgcc tgacaacagg cagcgtggc 3780
	attgtggca ggatcatctt gtccggaaag cggccatca ttcccgacag ggaagtcctt 3840
	tacccggagt tcgatgagat ggaagagtgc gcctcacacc tcoccttacat cgaacaggga 3900
	atgcagctcg ccgaacaata caaacagaag gcaatcggtt tcgtgcacaaac agccaccaag 3960
	caagcggagg ctgtctcc cgtgggtggaa tccaagtggc ggaccctcga agccttctgg 4020
10	gccaagcata tggaaattt catcagcggg atacaattt tagcaggctt gtccactctg 4080
	cctggcaacc cccgatagc atcaactgtg gcattcacag cctctatcac cagcccgc 4140
	accacccaaac ataccctctt gtttaacate ctggggggat gggtgccgc ccaacttgc 4200
	cctcccaagcg ctgtttctgc ttctcgatgc gcccgcacatc tcgggc 4260
	ataggccttgc ggaagggtct tggatattt tggcagggtt atggagcagg gttggcag 4320
	gcgcctcggtt cctttaaggc catgagcggc gagatgcctt ccaccgagga cctggtaac 4380
15	ctactccctg ctatcccttc ccctggcgc ctagtcgtcg gggtcgtgtg cgcaagcgata 4440
	ctgcgtcgcc acgtggggcc aggaggagggg gctgtgcagt ggatgaaccg gctgatagcg 4500
	ttcgcgtcgcc ggggttaacca cgtctccccc aecgactatg tgctgtggag cgacgctgca 4560
	gcacgtgtca cttagatctt cttagatctt accatcaactc agctgtgaa gaggcttca 4620
	cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680
	gattggatat gcacgggtt gactgattt aagacctggc tccagtc 4740
	cgattggccg gagggtttttt ctctcatgt caacgtgggtt acaaggaggat ctggcggggc 4800
	gacggccatca tgcaaaaccac ctgcggccatc ggaggcacaga taccgggaca tggaaaaaac 4860
	ggttccatga ggatcggttgc gccttagggc tggatgatca cgtggcatgg aacattcccc 4920
20	attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaatta ttctagggcg 4980
	ctgtggcgcc tggctgtca gggatcgtg gaggttacgc gggtggggaa ttccactac 5040
	gtgacggggca tgaccactga caacgttaaag tgcccggtgc agtttccggc ccccgaaattc 5100
	ttcacagaagat tggatgggttgcgggttgc acgtacgtc cagcgtgaa accccctcta 5160
	cgggaggaggatc acatttctt ctggggggctc aatacattacc tggtgggtc acagctccca 5220
	tgcgagcccg aaccggacgt aecgtgtc acettccatgc taccggacc cttccacatt 5280
25	acggcggaga cggctaagcg taggtggcc agggatctc cccctctt ggcagctca 5340
	tcagctagec agctgtctgc gccttcctt aaggcaacat gcaactaccg tcactgactcc 5400
	cgggacgctg acctcatcgca ggccaaacctc ctggggggc aggatgtgg cgggaaacatc 5460
	accgcgtgg agtcagaaaa taaggtagta atttttgact cttagcggcc gctccaagcg 5520
	gaggaggatg agagggaaat atccgttccg gggagatcc tgccggggc cagggaaattc 5580
	cctcgagcga tgcccatatc ggcacggccg gattacaacc ctccactgtt agagtccctgg 5640
	aaggacccgg actacgtccc tccagtgta cacgggtgtc cttccggcc tggcaaggcc 5700
30	cctccgatatac caccctccacg gggaaaggagg acgggtgtcc tcgtcagaatc taccgtgtc 5760
	tctgccttgc cggagctcgc cacaagacc ttgcgcagct cggaaatcg 5820
	agcgccacgg caacggccctc ttctgaccat ccctccgacg acggcgacgc gggatccgac 5880
	gttgagtcgtt accttcatc gccccccctt gaggggggatc cggggggatcc egatctcagc 5940
	gacgggttctt ggttacccgt aacggggagg gctatgtggg agtgcgtctg ctgtctcgat 6000
	tccatcacatc ggacaggcgc ctgtatcgc ccatcggtc cggaggaaac caagctgccc 6060
	atcaatgcac tgagcaactc tttgtccgtt caccacaact tggttatgc tacaacatct 6120
35	cgcagcgcac gctcgccca gaagaagggtt accttgcaca gactgcaggat cctggacgac 6180
	cactaccggg acgtgtctaa ggagatggaa gggaggcgatccacatgtttaaaatttgc 6240
	ctatccgtgg aggaaggcttca taatgtacg ccccccattt cggccatggc taaatttgc 6300
	tatggggcaaa aggacgtccg gaacccatcc acggatcccg ttaaccatcc cgcgtccgt 6360
	tggaaaggact tgctggaaat cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
	gaggttttctt gcttccaaacc agagaagggg gggcccaagec cagctcgctt tategtattc 6480
40	ccagatttgg ggggttcgtt gtgcggaaaaa atggccctt acgtatgtgtt ctccaccctc 6540
	cctcaggccg tggatggctt ttcataatcgatc ttccaaactt ctccgggaca ggggggtcgag 6600
	ttctctggta atgcctggaa agcggaaaaa tgccctatgg gcttcgcata tggatcccg 6660
	tgttttactt caacggtcac tgagaatgtac atccgtgttgg agggatcaat ctaccaatgt 6720
	tgtgacttgg cccccggaaac cagacaggcc ataagggtcgc taccaggacg gctttacatc 6780
	ggggggccccc tgacttaatcc taaaggggatc aactcgccgtt atccgggtt cccggcgagc 6840
45	ggtgtactga cggaccatgtt cggtaatacc ctccatgtt atctgtggc cgtcgccggc 6900
	tgtcgagctg cgaagctca gggactgcacg atgtcgatc gggagacga ctttgtcggtt 6960
	atctgtgaaa gggggggggcc ccaagaggac gggggggcc tacggggctt cacggggatc 7020
	atgactagat actctgcccc ccctggggac ccggccaaac cagaatacga cttggatgtt 7080
	ataacatcatc gtcctccaaat tggatcgatc gggccacatc catctggcaa aagggtgtac 7140
	tatctccatcc ttgtacccatc cacccttccatc ggggggttgc ctttgtcgatc 7200
	actccagtcac attcctggctt aggacacatc atcatgtatc cggccacatc ttggggatcc 7260
50	atgatcttcgtt tgacttcattt ctctccatc ttcttagctc aggaacactt tggaaaaagcc 7320
	ctagattgtc agatctacgg ggccttttac tccattggatc cacttgacatc acctcagatc 7380
	attcaacgcac tccatggctt tagcgcattt tcaactccata tttactctcc aggtgagatc 7440
	aatagggtgg cttecatgtc caggaaactt ggggttccatc ctttgtcgatc 7500

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
tcccagttgg atttatccag ctggtcgtt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccggacc ccgctggttc atgtggtgcc tactctact ttctgttaggg 7740
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca gccaaatagg 7800
ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttctt ctttttttcc ttttcttcc tttgggtggct ccatcttagc 7920
cctagtcacg gctagctgtg aaaggccgt gagecgctt actgcagaga gtgctgatac 7980
10 tggcctctct gcaagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 4
<211> 7989
<212> DNA
<213> Hepatitis C Virus

5	<400> 4
10	gcccagcccc gattgggggc gacactccac catagatcac tcccccgtgta ggaactactg 60 tttcacgcga gaaagcgctc agccatggcg ttatgtatgat tgctgtgcag cctccaggac 120 ccccccccc gggagagcca tagtggtctg cggAACCGGT gtagtacaccg gaattgcac 180 gacgaccggg tcctttctt gatcaacccg ctcacatgcct ggagatggg gctgtcccc 240 gcgagactgc tagccgagta gtgttgggc gcggAAAGGC ttgtgttact gcctgtatagg 300 gtgttgcga gtggccccc aggtctgtc gaccgtgcac catgagacac aatcttaaac 360 ctaaagaaa aaccaaaggg cgcccatga ttgaaacaatg tggatgtc acgggttctc 420 cgccgcgtt ggtggaggg ctatggct atgactggc aacacagaca atcggtctc 480 ctgtgcgcg cgtgttccgg ctgtcagcgc agggcgccc gtttctttt gtcaagaccc 540 acctgtccgg tgccctgaat gaactgcagg acgaggcgc gcggctatcg tggctggca 600 cgacggcggt tccttgcgcg gctgtgtcg acgttgtc ac tgaaggcgga aggactgtgc 660 tgetattggg cgaatgtcg gggcaggatc tcttgtcata tcaccccttgc cctgccgaga 720 aagtatccat catggctgt gcaatgcgc ggctgcatac gttgtatcg gctacactgc 780 catcgacca ccaagcgaaa catcgcatcg akgcagcag tactcgatg gaagccggtc 840 ttgtcgatca ggatgtatcg gacgaagac atcaggggct cgcgcacagcc gaactgttcg 900 ccagcgctaa ggcgcgcgt cccgcgcgg agatctgt ctgtacccatgc ggcgtatgc 960 gcttgcggca tatcatgggt gaaatggc gctttctgg atcatcgac tggccggc 1020 tgggtgtggc ggaccgtat caggacatag ctgtggctac ctgtgtatatt gctgtacag 1080 ttggcgccga atgggtgtac cgcttctcg tgctttacgg tattcgccgt cccgatctgc 1140 agcgcatcgc cttctatcgc cttctgtac agtttttctg agtttaaaca gaccacaacg 1200 gtttccctct agcgggatca attccgcggc tctccctccc ccccccctaa ctgtactgtc 1260 cgaagccgcgt tggaaatagg cgggtgtcg tttgttata ttttttttcc caccatattg 1320 cctgttttgc gcaatgtcg gggccggaaa ctggccctg ttttttttgc gaggatctt 1380 aggggttctt ccctctcgc caaaggatg caaggctgtg tgaatgtcgta gaaggaagca 1440 gttctcttgg aagtttctg aagacaaaaca acgtctgtag cgaccctttg caggcagccg 1500 aacccccacat ctggcgacag gtgcctctgc ggcggaaacg cactgtatca agatacacct 1560 gcaaaaaaccat cacaatccatgt gtggccatgt gtgagttgg tagttgttga aaggtcaaa 1620 tggcttcttgc caagcgat ttttttttgc ctttttttgc ttttttttgc ttttttttgc 1680 atggatctg atctggggcc tcgggtcaca tgctttacat gtgttttagtc gaggttaaaa 1740 aacgtcttagg ccccccgaac cacggggacg tgggttttctt tgaaaaaaca cgataatacc 1800 atggcgccatc ttacggctca ctcccaacag acgcggaccc tactttgtcg catcatact 1860 agcctcacag gcccggacag gaaccaggc tgggggggg tccaaatgtgttgc 1920 acacaatctt tcttgcgcgc ctgcgtcaat ggctgtgtt ggactgtctt tcatgtgtcc 1980 ggctcaaaa cccttgcgg cccaaaggc ccaatcaccc aaatgtacac caatgtggac 2040 caggacctcg tcggctggca agcgcccccc gggcgccgtt ctttgacacc atgcacctgc 2100 ggcagctcggt accttttactt ggtcagggc catggcgtatg ttttttttgc tggccggccg 2160 ggcgcacagca gggggaggt acttcccccc agggccgtt ctacttgaa gggcttctcg 2220 ggcggttccat tgctctggcc ctggggcact gttgtggca ttttttttgc tggctgtgc 2280 acccgggggg ttgcgaaggc ggtggacttt gtacccgtcg agtctatggta aaccactatg 2340 cggtccccgg tcttcacgga caactcgcc cctccggccg taccgcagac attccaggtg 2400 gccccatctac acgcggccatc tggtatggc aagagacta aggtggccggc tgctgtatgc 2460 gccaagggtt ataagggtt tgctcttgcac ccgtccgtcg ccggccaccct aggtttccgg 2520 gcgtatatgt ctaaggcaca tggtatcgac cttaaatctca gaaccgggtt aaggacatc 2580 accaagggtt ccccatctac gtactccacc tatggcaagt tttttgcgcg cgggtgtgc 2640 tctggggcgct cctatgacat cataatatgt gatgagtgcc actcaactga ctgcaccact 2700 atccctggca tcggcacatg cctggaccaat gggagacccg ctggagcgcg actctgtctg 2760 ctcgccacccat ctacggctcc gggatcggtc accgtgcacat cttaaatctca cgaggagggt 2820 gctgttgcac gcaactggaa atccccctt ttttttttgc aatcccccattt cggacatc 2880 aagggggggg ggcacccat ttttttttgc tccaaagaaga aatgtatgt gtttttttgc 2940 aagctgtccg gcttcggact caatgtcgta gcatattacc gggcccttgc ttttttttgc 3000 ataccaacta gcccggacgt cattgtcgta gcaacggacg ctctaatgtc gggctttaacc 3060 gcccgttccg actcagtgtatcgactcgatcatgtgc ttttttttgc cccaggacat cggacttgc 3120 ctggacccggat ctttccatcat tgagacggacg accgtgcacat aacggccgtt gtcacgtcg 3180 cagcggccggat gcaaggactgg tagggcagg atgggcattt acagggtttgt gactccagg 3240 gaacggccctt cggggatgtt cgatttctcg gttttgtcg agtctatgtc cggggctgt 3300 gcttggtaac gcttcggact cggccggaccc tcaatgtgttgc ttttttttgc cttaaacaca 3360 ccagggttgc cctgttgcac gggatctgtc gggatctgttgc agggccgtt tacaggccctc 3420 acccacatag acggccatcat ttttttttgc actaaggcagg cggaggacaa ctttttttgc 3480 ctggtagcat accaggatcat ggtgtgcgcg agggtgcagg ttttttttgc atctgtggac 3540

caaatgtgga agtgttccat acggctaaag cctacgctgc acggggcaac gcccctgctg 3600
 tataggctgg gagccgttca aaacgagggtt actaccacac accccataac caaatacattc 3660
 atggcatgca tgcgttgc cctggagggtc gtcacgagca cttgggtgtc ggttagcgga 3720
 gtccatcgag ctctgcgcgt gtattgtctg aacaaggca gcgtgttcat tgcgttgcagg 3780
 atcatcttgc cggaaagcc ggcatttgc cccgacaggaa aagtccattt ccgggagttc 3840
 gatgagatgg aagagtgcgc ctcacacccctt ctttacatcg aacagggttgc gcaagctcgcc 3900
 gaacaaatcca aacagaaggc aatcggttgc ctttacatcg aacagggttgc gcaagctcgcc 3960
 gctgtccccc tggatgttgc ctttacatcg aacagggttgc gcaagctcgcc 4020
 tggatgttca tcagcggat acaatatttgc ctttacatcg aacagggttgc gcaagctcgcc 4080
 gcgatagcat cactgttgc attcacagcc ttttacatcg aacagggttgc gcaagctcgcc 4140
 acccttcctt ttaacatccctt gggggatgg gtggccccc aacttgcctt tcccttcgtt 4200
 gcttctgtt tcgttgcgc cggatcgatc gggatcgatc ttggcagcatc aggccttggg 4260
 aaggatgttgc tggatgttgc ctttacatcg aacagggttgc gcaagctcgcc 4320
 ttaacatcgatc ttttacatcg aacagggttgc gcaagctcgcc 4380
 atcccttcctt ctggccccc ttttacatcg aacagggttgc gcaagctcgcc 4440
 gtggggccatc gggagggggc ttttacatcg aacagggttgc gcaagctcgcc 4500
 ggttacccatc ttttacatcg aacagggttgc gcaagctcgcc 4560
 cagatcttctt ctgttgcgc ctttacatcg aacagggttgc gcaagctcgcc 4620
 gaggactgttgc ctttacatcg aacagggttgc gcaagctcgcc 4680
 acggatgttgc ttttacatcg aacagggttgc gcaagctcgcc 4740
 gtttttttttctt ttttacatcg aacagggttgc gcaagctcgcc 4800
 caaacatccatc ttttacatcg aacagggttgc gcaagctcgcc 4860
 atcgatgggc ttttacatcg aacagggttgc gcaagctcgcc 4920
 accacggggcc ctttacatcg aacagggttgc gcaagctcgcc 4980
 gctgttgcagg agtacgttgc ggttacggggttgc gcaagctcgcc 5040
 accactgttgc aacgttgcgttgc gtttccggccatc ctttacatcg aacagggttgc gcaagctcgcc 5100
 gatgggggtgc ggttacggggttgc gcaagctcgcc 5160
 acatcttgc ttttacatcg aacagggttgc gcaagctcgcc 5220
 ctttacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 5280
 gtaaaggatc ggttacggggttgc ggttacggggttgc gcaagctcgcc 5340
 ctgttgcgc ttttacatcg aacagggttgc gcaagctcgcc 5400
 ctatcgagg ctttacatcg aacagggttgc gcaagctcgcc 5460
 ttttacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 5520
 agggaaatcgatc ttttacatcg aacagggttgc gcaagctcgcc 5580
 ctttacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 5640
 taatcgatc ttttacatcg aacagggttgc gcaagctcgcc 5700
 ctttacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 5760
 gagatgttgc ttttacatcg aacagggttgc gcaagctcgcc 5820
 acggccatc ttttacatcg aacagggttgc gcaagctcgcc 5880
 ttttacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 5940
 tetacatcg aacagggttgc ggttacggggttgc gcaagctcgcc 6000
 acaggccatc ttttacatcg aacagggttgc gcaagctcgcc 6060
 agcaactttt ttttacatcg aacagggttgc gcaagctcgcc 6120
 ctggccatc ttttacatcg aacagggttgc gcaagctcgcc 6180
 gtgttgcgc ttttacatcg aacagggttgc gcaagctcgcc 6240
 gaaggcttgc ttttacatcg aacagggttgc gcaagctcgcc 6300
 gacgttgcgc ttttacatcg aacagggttgc gcaagctcgcc 6360
 ctggaaatcgatc ttttacatcg aacagggttgc gcaagctcgcc 6420
 gtccaaatcgatc ttttacatcg aacagggttgc gcaagctcgcc 6480
 gtggatgttgc ttttacatcg aacagggttgc gcaagctcgcc 6540
 atgggttgc ttttacatcg aacagggttgc gcaagctcgcc 6600
 gctgttgc ttttacatcg aacagggttgc gcaagctcgcc 6660
 acggatgttgc ttttacatcg aacagggttgc gcaagctcgcc 6720
 ctttacatcg aacagggttgc gcaagctcgcc 6780
 actaatttc ttttacatcg aacagggttgc gcaagctcgcc 6840
 accatgttgc ttttacatcg aacagggttgc gcaagctcgcc 6900
 aagcttgc ttttacatcg aacagggttgc gcaagctcgcc 6960
 gggggatggc ttttacatcg aacagggttgc gcaagctcgcc 7020
 ttttacatcg aacagggttgc gcaagctcgcc 7080
 ttttacatcg aacagggttgc gcaagctcgcc 7140
 gacccatc ttttacatcg aacagggttgc gcaagctcgcc 7200
 ttttacatcg aacagggttgc gcaagctcgcc 7260
 actcattttt ttttacatcg aacagggttgc gcaagctcgcc 7320
 atctacggg ttttacatcg aacagggttgc gcaagctcgcc 7380
 catggccatc ttttacatcg aacagggttgc gcaagctcgcc 7440
 teatgttgc ttttacatcg aacagggttgc gcaagctcgcc 7500

EP 1 043 399 A2

5
gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 7560
aactggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 7620
ttatccagct ggttcggtgc tggttacagc gggggagaca tatatcacag cctgtctcg 7680
gcccgcaccc gctgggtcat gtgggtgccta ctcctacttt ctgttaggggt aggcatctat 7740
ctactccccca accgatgaac ggggagctaa acactccagg ccaataggcc atccctgtttt 7800
tttccctttt tttttttctt tttttttttt tttttttttt ttctccctttt 7860
tttttcctct ttttttcctt ttctttcctt tggtggctcc atcttagccc tagtcacggc 7920
tagctgtgaa aggtccgtga gccgcttgac tgcaagagat gctgatactg gcctctctgc 7980
10 agatcaagt 7989

15

20

25

30

35

40

45

50

55

<210> 5
 <211> 8649
 <212> DNA
 <213> Hepatitis C Virus
 5
 <400> 5
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagctct agccatggcg ttagtatgag tgcgtgcag cttccaggac 120
 ccccccctccc gggagagcca tagtggctg cggAACCGGT gagtacaccg gaatttccag 180
 gacgaccggg tccttcttg gatcaacccg ctcaatgcct ggagatttg gcgtgcccc 240
 10 gcgagactgc tagccgatg gtgtggtc gcgaaaggcc ttgtggact gcctgatagg 300
 gtgttcgca gtccgggg aggtctcgta gaccgtgcac catgagcag aatcccaaac 360
 ctcaaaagaa aaccaaactg aacccaacg ggcgcgcac gattgaacaa gatggattgc 420
 acgcagggtc tccggccgct tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
 caatcggtc ctctgtatgcc gccgttcc ggctgtcagc gcaggggcgc ccggttctt 540
 ttgtcaagac cgaccgtc ggtgcctga atgaactgc ggacgaggca ggcgcgtat 600
 cgtgttcgca caccggggc gttccttcg cagctgtgtc cgacgttgc actgaagcgg 660
 15 gaaggggactg gtcgtatcc ggcgaatgtc cggggcaggaa tcttcgttca ttcacccctt 720
 ctccgtccca gaaatgtatcc atcatgtgtc atgcaatgcg gggcgtcat acgtttgatc 780
 cggttacccg cccattcgac caccaggcg aacatcgcat cgagcgagca cgtactcgga 840
 tggaaaggccg tcttgtcgtat caggatgtc tggacgaaaga gcatcagggg ctcgcgcag 900
 cccaaactgt cgcggcgtc aaggcgcgc tggccgcacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttcgtc aataatcgatgg tggaaaatgg ccgttttctt ggtttatcg 1020
 20 actgtggccg gtcgggtgtc gcgaccgcgt atcaggatc agcgttgc acccggtata 1080
 ttgtcaaga gcttggccg gaatgggtc accgttctt cgtgttttac ggtatcgccg 1140
 ctcccgattc gcagcgatc gccttcatttgc ctgccttgcg ctagtttttc tgagttaaa 1200
 cagaccacaa cggtttccct ctacgggtat catttcggcc ccttccttcc cccccccctt 1260
 aacgttactg gccgaaggcg ctggaaataa ggccgggtgtc cgtttgtata tatgttattt 1320
 tccaccatat tgccgttcc tggcaatgtc aggccccggaa aacctggccc tgcgttcc 1380
 25 acgagcatc ctagggttct ttcccttcg cccaaaggaa tgcaagggtc gttgaatgtc 1440
 gtgaaggaag cagtttctt ggaagcttca tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcgc ggaacccccc acctggcgc acgtgttgc gggccaaaa gccacgtgt 1560
 taagataacat ctgcaaggc ggcacaccc cagtgcacg ttgtgatgg gatagttgt 1620
 gaaagagtc aatggcttc ctaaagcgta ttcaacaagg ggctgaagga tggccagaag 1680
 gtacccatt gtatggatc tgatctggg ctcgggtcga catgttttac atgtgttag 1740
 30 tcgaggtaa aaaacgtcta ggcggcccg accacggggc cgtgttttc ctttggaaaa 1800
 cacgataata ccatggaccg ggatggca gcatcggtcg gggcgggt ttctgttaggt 1860
 ctgtacttct tgacccgtc accgcactat aagctgttc tgcgttagt catatggtg 1920
 ttacaattat ttatcaccag ggcggaggca cacttgcacg tggatccc cccctcaac 1980
 gttcgggggg gccgcgtac cgttccatc ctcacgtcgc cgatccaccc agagctaata 2040
 ttaccatca cccaaatctt gctccatca ctccgggttca tcatgggtc ccaggcttgt 2100
 ataaccaaaatg tggccgtactt cgtgcgcga caccgggtca ttcgtgcgt catgtgttg 2160
 35 cggagggttgc ctgggggtca ttatgtccaa atggctcttca tgaagggttgc cgcactgaca 2220
 ggtacgtacg ttatgtacca tctccatcca ctccggact gggccacgc gggccatcga 2280
 gaccttgcgg tggcagttga gcccgtcgtc ttctctgata tggagacaa ggttatcacc 2340
 tggggggcag acaccgcgc gttggggac atcatctgg gcctgcccgt ctccggccgc 2400
 agggggaggc agataactat gggacccggc gacagccgttgc aaggccagg gtggcactc 2460
 ctccgcctttaatggccta ctcccaacag acgcggaggcc tacttgcgt catcatca 2520
 40 agcctcacac gccggggacac gaaaccaggc gggggggagg tccaaatggt ctccaccgc 2580
 acacaatctt tcctggcgac ctgcgtcaat ggctgtgtt ggactgtcta tcatgggtcc 2640
 ggctcaaaga cccttgcgg cccaaaggcc ccaatcaccc aaatgtacac caatgtggac 2700
 caggacactcg tcgggtgc acgcggccccc ggggcgcgt ccttgcaccc atgcacactgc 2760
 ggcagctcg acctttactt ggtcaegagg catgcgtatgc tccatccgtt ggcggccgg 2820
 ggcgacacca gggggggccctt actctccccc aggccccgttgc tccatcttgc 2880
 ggcgggtccac tgctctgtccctt ctccgggtcgc gctgtggca tcttcgggc tgcgtgtgc 2940
 45 acccgagggg ttgcgaaggc ggtggacttt gtacccgtcg agtctatgg aaccactatg 3000
 cggccccccg tcttcacgca caactcgcc ctcggccgcg taccgcagac attccagggt 3060
 gccccatctac acgcggccctt tgtagccgc aagagcacta aggtgcggc tgcgtatgc 3120
 gcccgggttataagggtct tgcgtatgcac ccgtccgtcg ccgcacccat aggttcggg 3180
 qcgatataatgt ctaaaggcaca tggatcgac cctaacaatca gaaaccgggtt aaggaccatc 3240
 accacgggttgc ccccaatcac gtaactccacc tatggcaagt ttcttgcga cgggtgttc 3300
 50 tctggggggcctatgtatgt gatgagtgcc actcaactga ctcgaccact 3360
 atcttggggca tcggcactgt cctggacca gggggacccgg ctggagcgcgc actcggtcg 3420
 ctccggccaccctt ctacggccccc gggatcggtc accgtgcaccc atccaaacat cgaggagggt 3480
 gtcgttccca gcaactggaga aatccctt tattggcaag ccattccatc cgagaccatc 3540

5	aaggggggga ggcacccat tttctgccc tccaagaaga aatgtatgt gctcgccg 3600 aagctgtccg gcctcgact caatgtgtt gcatattacc ggggccttga tgtatccgtc 3660 ataccaacta gcggagacgt cattgtgtt gcaacggacy ctctaatacg gggctttacc 3720 ggcgttccg actcagtgtt gcaactgtt acatgtgtt cccagacagt cgaacttcacg 3780 ctggaccctt ccttccatc tgagacgacg accgtgccac aagacgcggt gtcacgtctc 3840 cagcgccgag gcaggactgg tagggcagg atgggcattt acaggtttgt gactccagga 3900 gaacggccctt cgggcattt cgatttctcg gttctgtcg agtgcatacg cgccggctgt 3960 gcttggtaacg agtcacgcg cggcggacg tcgatgttgc tgccggctta cttaaacaca 4020 ccagggttc ccgtctgca ggacatctg gaggctctgg agagcgtctt tacaggcctt 4080 acccacatag acgccccattt cttgtccccag actaagcagg caggagacaa cttcccttac 4140 ctggtagcat accaggctac ggtgtgcgc agggtcagg ctccacacttcc atcggtggac 4200 caaatagtgg aagtgtctcat acggctaaag cttacgcgtc acggggcaac gcccctgtcg 4260 tataggctgg gaggcgatc aaacggatgtt actaccacac accccataca caaatacatc 4320 atggcatgcg tggcgttgta cttggggatc gtacggatc cttgggtgtc ggtaggccga 4380 gtccctgtcg ctctggccgc ttatgtctgtt acacaggca ggtgtgtcat tggtggccgg 4440 atcatcttgc cccggaaagcc ggccatcattt cccgacaggg aagtccctta cccggagggtc 4500 gatgagatgg aagagtgcgc ctcacacccctt cttacatcg aacaggaaat gcaagctcgcc 4560 gaacaatcc aacagaaggc aatcgggtt ctgcacaaacg ccaccaacgca agccggaggct 4620 gctgtcccg tggtggatc caagtggcg accctcgaaat cttctggc gaagcatatg 4680 tggaaatttca tcagcggtt acaaatattt gagggttgc ttactctgtc tggcaaccc 4740 gcgatagcat cactgtatgc attcacagcc ttatcacca gcccgtcactt caccacacat 4800 accctctgtt ttaacatctt gggggatgg gtggccgccc aacttgcctt cccagegtct 4860 gcttctgtt tcgttaggcgc cggcatcgct ggacggctgt ttggcagatc agcccttggg 4920 aagggtcttgg tggatattttt ggcaggatgtt ggacggaggcc tggcaggcgc gctgtggcc 4980 ttaaaggctca tgagcgccgc gatgccttcc accggaggacc ttgtttaacctt atccctgtt 5040 atccctcttcc ctggcgccctt agtgcgtggg gtgtgtgtcg cagcgataact ggtcgccac 5100 gtggggccag gggagggggc tggcgtatgg atgaaccggc tgatagcggtt cgcttcgccc 5160 gttaaccacg tctcccccac gcaactatgtt cttggagagcc acgtgtcgac acgtgtact 5220 cagatccctt ctatgtttt catactctg ctgtcaagaat ggttccatca gttggatcaac 5280 gaggactgtt ccacgcctatg tcccggtctg ttggtaagag atgtttgggaa ttggatatgc 5340 acgggtttgtt ctgatttcaaa gacgggtctt cttgttccacttcc gttgttccatc gtcgtccatcg 5400 gttcccttctt tctcatgtca acgtgggtac aagggtgtt ggcggggccg cggcatcatg 5460 caaaccacat gcccattgtt gacgacatcg accggatcatg tgaaaacaccc ttccatgagg 5520 atcgtggggc cttaggactt tagtaacacg tggcatggaa catttccatc taacgcgtatc 5580 accacggggc cttgcacccctt cttcccgccg ccaattttt ctatggcgtt tggcgggtgt 5640 gctgtgttgg agtacgtgtt ggttacggcg gtgggggattt tccactacgt gacgggtatg 5700 accactgaca acgtaaatgtt cccgtgttccg gttccggccc cggaaatttccatc cacaagaatg 5760 gtatgggtgc ggttgcacatc gtacgttccca gctgtcaac ccccttcttacatc ggagggtgtc 5820 acattcttgc tgggcttcaaa tcaatactgtt gttgggttccatc agtcccccattt cggccggaa 5880 ccggacgtatc cttgtcttcc ttccatgtctt accggaccctt cccacattac ggcggagac 5940 gttaaggctt ggttggccatg gggatcttcc ccccttcttg ccagtcatac agctagccag 6000 ctgtctgtcc cttcccttgc ggcacatcgacttcccttccatc gtcgttccatc ggacgtgtac 6060 cttcatcgatc ccaacccttctt gttggccgcg gagatggccg gaaacatcactt cccgtgtgg 6120 tcagaaaaata aggttaatgtt ttggatctt ttcggatccatc tcaacggca ggaggatgtt 6180 agggaaatgtt ccgttccggc ggagatctgtt cggaggttca gggaaatttccatc tccggatgtc 6240 cccatatggg caccggccgc ttacaaaccctt ccactgtttag agtctggaa ggaccggac 6300 tacgtcccttcc cttgtgttca cgggtgttccca ttggccgttcc ccaaggcccccc tccgatacc 6360 cttcacccggat ggaagaggac ggttgcgttcc tcaatactt cttgtgttccatc tgcccttggc 6420 gagctcgccat caaacactttt cggcgttcc gatctgttcc cctgtcgacag cggcaccgg 6480 acggcccttc ctgaccacccctt cttccgttcc ggcgttccatc gatccgtatc tggatgttcc 6540 ttctccatgtt ccccccatttgc gggggggccg ggggatcccg atctcacttgc cgggttcttgg 6600 tctaccgtttt gcgaggaggc tagtggaggac gtcgttctgtt gtcgtatgttcttccatc tcaacatcg 6660 acaggcccttc tgatcactccatc atgcgttccg gggaaacccatc gtcgttccatc gtcgttccatc 6720 agcaactttt tgcctgttccatc ccacacttgc gttctgttca ccaatctcg cagcgcaac 6780 ctggccggatc agaagggttccatc ctttgcacccatc tggcggatccatc tggacgacca 6840 gtgtctcaagg agatgtttttgc gggggggccg ggggatcccg atctcacttgc cgggttcttgg 6900 gaaggccgttgc agctgtatccatc cccacatttcg gccagatcttca aatttggatccatc tggggccaaag 6960 gacgtccggatc acctatccatc caaggccgttccatc aaccacatccatc gtcgttccatc gggaggactt 7020 ctggaaagaca ctggacacccatc aatttgcacccatc tggatgttccatc cccatctcg tggatgttccatc 7080 gttccaaaccatc agaaggatccatc ctttgcacccatc gatgtgttccatc tggatgttccatc tggatgttccatc 7140 gttccgttgc gcgagaaaaat gggcccttccatc gatgtgttccatc cccacccatc tccaggccgt 7200 atgggtctttt cttttttccatc ccaatacttgc cttggacatccatc ggggtgttccatc tggatgttccatc 7260 ggcttggaaatc cttttttccatc cccatcttgc gttctgttccatc aaccacatccatc tggatgttccatc 7320 acggtctacttccatc agaaggatccatc ctttgcacccatc gatgtgttccatc tggatgttccatc tggatgttccatc 7380 cccgaaagccatc gacaggccatc aagggttccatc acagacggccatc tttatctcg gggcccccgt 7440 actaatttcttca aagggttccatc ctggccgtatc cggccgttccatc gggcccccgt 7500
10	
15	
20	
25	
30	
35	
40	
45	
50	

accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgc 7560
 aagctccagg actgcacgt gctcgatgc ggagacgacc ttgtcggtat ctgtgaaagc 7620
 5 gcggggaccc aagaggacga ggcgagccta cgggcctca cggaggtat gactagatac 7680
 tctgcccccc ctggggaccc gcccaaacca gaatacgtact tggagggtat aacatcatgc 7740
 tcctccaatg tgtcagtgcg gcacgatgca tctggcaaaa gggtgtacta ttcacccgt 7800
 gacccacca ccccccattgc ggggctgcg tgggagacag ctagacacac tccagtcaat 7860
 10 tcctggctag gcaacatcat catgtatgcg cccaccttg gggcaaggat gatcctgtat 7920
 actcatttct tctccatctt tctagctcg gaacaacttg aaaaaggctt agattgtcg 7980
 atctacgggg cctgttactt cattgagcca cttgacctac ctcagatcat tcaacgactc 8040
 catggccta ggcgcatttc actccatagt tactctccag gtgagatcaa tagggtggt 8100
 15 tcatgccta gaaaaactgg ggtaccgccc ttgcgagtct ggagacatcg ggccagaagt 8160
 gtcccgctt ggcgtactgtc ccaggggggg agggtgccta cttgtggcaa gtacctttc 8220
 aactggcag taaggaccaa gtc当地actc actccatcc cggctcgctc ccagttggat 8280
 ttatccagct gggtcggtgc tggttacagc gggggagaca tatatcacag cctgtctcg 8340
 gcccacccc gctggttcat gtgggccta ctccacttt ctgttaggggt aggcatctat 8400
 ctactccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgttt 8460
 tttccctttt tttttttttt tttttttttt tttttttttt ttcteetttt 8520
 tttttccctt tttttccctt ttcttcctt tgggtggctcc atcttagccc tagtcacggc 8580
 tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 8640
 20 agatcaagt 8649

25

30

35

40

45

50

55

<210> 6
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5 <400> 6
 gccagcccc gattggggc gacactcac catagatcac tcccctgtga ggaactactg 60
 ttttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cttccaggac 120
 ccccccctccc gggagagcca tagtggctg cgaaaccgg gatgacaccg gaattgcag 180
 gacgaccggg tcctttctg gatcaacccg ctcaatgcct ggagatttg gcgtcccc 240
 gcgagactgc tagccgaga gtgttgggtc gcaaggcc ttgtgtact gcctgatagg 300
 gtgcgtcgca gtggccccc aggtctcgta gacgtgcac catgacccg aatccaaac 360
 ctcaaaagaaa aaccaaaacg aaccaacg ggccgcctat gatgaaaca gatggattgc 420
 acgcagggtc tccggccgt tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtttcc ggctgtcage gcaggggcgc ccggttctt 540
 ttgtcaagac cgacctgtcc ggtgcctga atgactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacggc gttcctcgta cagctgtct gcacgtgtc actgaagcgg 660
 gaaggactg gctgttattt ggcgaatgtc cggggcagg tctctgtca ttcacccctt 720
 ctctgcgcga gaaatatacc atcatgctg atgacatgcg ggggtgcac acgcttgc 780
 cgcttacctg cccatcgac caccaacgca aacatcgcat cgacgagca cgtactcg 840
 tgaagccgg tcttgcgtc caggatgatc tggacaaga gcatcagggg ctcgcgcag 900
 ccgaactgtt cgccagggtc aacgcgcgc tggccgcgg cggaggatctc gtcgtgaccc 960
 atggcgatgc tctgtgcg aatatacatgg tggaaaatgg cgcctttctt ggattcateg 1020
 actgtggccg gtcgggtgtc gccggacccgt atcaggacat acgttgtggct acccgtata 1080
 ttgtcaaga gtttgcggc gaatgggtc accgtttctt cgtgtttac ggtatcgcc 1140
 ctccccgttgc acagcgcate gccttcatac gccttcttgc cagtttttc tgagttaaaa 1200
 cagaccacaa cggttttccctt ctagggatc caatccgcctt cctctccctt cccccccctt 1260
 aacgttactg gccgaagccg cttggataaa ggcgggtgtc cgtttgtca tatgttattt 1320
 tccaccatatac tccgtcttgc tggcaatgtg aggccggaa aacatggccc tgccttttgc 1380
 acgagcattc ctagggttctt tcccccttc gccaaaggaa tgcagggtct gttgaatgtc 1440
 gtgaaggaaag cagttccttgc ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcgcg ggaacccccc acctggcgac aggtgcctct gcccggaaaaa gccacgtgt 1560
 taagatacacatc ctgcaaaaggc ggcacaaaccc cagtgccacg tttgtgttgc gatagttgt 1620
 gaaagagtca aatggcttc tccaaaggca ttcaacaaagg ggctaaagg tgcggcggaa 1680
 gtacccattt gtagggatc ttagttttttt ctcgggtca catgttttac atgtgtttag 1740
 tcgaggttaaaa acacgtcta gggcccccga accacggggc cgtgttttc ctttggaaaa 1800
 cacgatataa ccatggccgca tattacggc tactccaaac agacgcggagg cctacttggc 1860
 tgcattatca ctacgttcac agggccggac aggaaccagg tccgggggg ggttcaatgt 1920
 gtcctccaccg caacacaaatc ttctctggcg acctgggtca atggcgatgt tttgactgtc 1980
 tatcatggtg cgggtctaaa gacccttgc ggcggaaagg gccaatcac ccaaatac 2040
 accaatgtgg accaggactt cgtcggttgc caagggcccc cggggggcg ttccttgaca 2100
 ccatgcacctt cggcgacgttcc ggaccccttac ttggtcacga ggcgtccca gtttgcgttcc 2160
 gtgcgcggc gggccggacag cggggggggc ctactcttcc cggggccgtt ctcctactt 2220
 aagggtctt cggcggttcc actgtctgc cccctggggc atgtgttggg cattttcgg 2280
 gtcggcggtt gcaacccgggg ggttgcgaag ggggtggact ttgttacccgt cgagtctat 2340
 gaaaccacta tgcggtcccc ggttttacg gacaactcg tccctccggc cgttccgg 2400
 acattccagg tggcccatatc acacgcctt actgttagcg gcaaggacac taagggtggc 2460
 gtcgtatgc tggcccaagg gtataagggtt ctgttgcata accccgttgc cggccacc 2520
 cttagtttcg gggcgatatac tgcataaggca catgtatgc acccttaacat cagaacgggg 2580
 gttagggacca tcaccacggg tgcctccatc acgtacttca cctatggca gtttcttgc 2640
 gacgggtgtt gtcgtttttt cgttccatgac atcataatat gtgtatgttgc ccactcaact 2700
 gactcgacca ctatcttgcg catggcaca gtcctggacc aacggggac ggttggacg 2760
 cgactcgatcg tgcgtccac acgtactcgatc cggggatcg tccatgtgc acatccaaac 2820
 atcgaggagg tggctctgtc cagactggaa gaaatccctt ttatggca agccatcccc 2880
 atcgcgacca tcaagggggg gaggccatctt atttcttgc attccaaagaa gaaatgtat 2940
 gagctcgccg cgaagctatac cggccctcgga ctcaatgtcg tagcatatta cggggccctt 3000
 gatgtatccg tcataccaac tagggggac gtcgttgc tgcgttgc gctcttaatg 3060
 acgggttta cggcgatattt cgtactgtt atcgactgtca atcatgttgc caccggac 3120
 gtcgacttca gtcgtggacc gacccgttcc accggatccg tccatgtgc acaagacgc 3180
 gtgtcacgtt cgcacggccg aggccggact ggttggggca ggttggccat ttacagggtt 3240
 gtgactccag gagaacggcc ctcggccatg ttgttccctt cgggttgcgtc cgagtgttat 3300
 gacgggggtt gtcgttggta cgtactgtc cccggggaga ccttgcgttgc gtttggggct 3360
 tacctaaaca caccggggctt gcccgttgc caggaccatc tggatgttgc ggagacgc 3420
 ttacaggcc tcaaccatc agacggccat ttcgttgc agactaggca ggcaggagac 3480
 aactccctt acctggtagc ataccaggctt acgtgtgcg ccagggttca ggctccac 3540

5 ccatcggtgg accaaatgtg ggagtgtctc atacggctaa agcctacgct gcacgggcca 3600
 acgccccctgc tttataggct gggagccgtt caaaacgagg ttactaccac acaccccata 3660
 accaaataca tcatgcgtat catgtcggtt gacctggagg tcgtcacag cacctgggtg 3720
 ctggtaggcg gagtcttagc agctctggcc gctgtattgcc tgacaacagg cagcgtggc 3780
 attgtggcga ggatcatctt gtccggaaag ccggccatca ttcccacag ggaagtccct 3840
 taccgggagt tcgatggat ggaagagtgc gcttcacacc tcccttacat cgaacaggga 3900
 atgcagctcg ccgaacaaatt caaacagaat gcaatcggt tgctgcacac agccaccaag 3960
 caaaggcggt ctgctgtcc ctgggtggaa tccaaatggc ggacccctcg agccttctgg 4020
 gcaagcata tttggaaattt catcagcggtt atacaatatt tagcaggctt gtccactctg 4080
 10 cctggcaacc ccgcgtatgc atcaactgtatg gatttcacag cctctatcac cagcccgctc 4140
 accacccaaat ataccctctt gtttaacatc tggggggat gggtgccgc ccaacttgc 4200
 cctccacgc tctgttcgc tttctgtgc gccggcatcg ctggagccgc tggtggcage 4260
 ataggccctt ggaagggtct tttggatattt tggtggatggt atggcggatc ggtggcaggc 4320
 ggcgtcggtt ctttaaggtt catgagcggtt gagatggccctt ccaccggaga cctggtaac 4380
 ctactccctg ctatcccttc ccctggcgcc ctatgtcgatg gggtcgtgtg cgcaggatc 4440
 ctgcgtcgcc acgtggggccc aggggggggg gctgtgcgtt ggtgtacccg gctgtatgc 4500
 ttgcgttcgc gggtaaccaat ctcagatctt ctctgttcc accatcactatg agtgcgtcaa gaggcttcac 4560
 gcacgtgtca ctcagatctt ctctgttcc accatcactatg agtgcgtcaa gaggcttcac 4620
 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggtaatc agatgtttgg 4680
 gattggatata gacgtgtt gactgttcc aagacgttgc tccagtcacaa gctccctggc 4740
 cgattggcg gagtccctt ctgttcgtat caacgtgggtt acaaggaggat ctggccggggc 4800
 gacggcatca tggcaaccat ctggccatgtt ggacccatggatc tccaggaca tggaaaaaac 4860
 gttccatca ggatcggtt ggcttggacc ttgttagtacaat cgtggatgg aacatcccc 4920
 20 attaacgcgtt acaccacggg ccctgtcactt ccctcccccgg cgcacaaatata ttctaggggc 4980
 ctgtggccggg tggctgttgc ggagtacgtg gagggttacgc ggggtggggg tttccactac 5040
 gtgacgggca tgaccactga caacgtaaatg tggccgtgtc aggttccggc ccccgaaattc 5100
 ttcacagaaat tggatggggt ggggttgcac aggtacgtcc cagcgtcaaa accccctctta 5160
 cggggggggg ttcacattctt ggtggggctt aatcaatacc tgggtgggtc acagctccca 5220
 tgcggggcccg aaccggacgtt agcgtgtcc acttccatgc tcacccggatc ctccacatt 5280
 acggggggaa cggctaaatcg taggtgtggcc aggggatctc cccctccctt ggcacgtca 5340
 25 tcagctatcg agtgcgttcgc ggcacccatc tttttccctt aaggacatc gcaactcccg tcatgactcc 5400
 cgggacgttgc acatccatcgca ggcacccatc ctgtggccgc aggagatggg cgggacacatc 5460
 accccgcgtgg agtcagaaaaaa taaggtagta atttttggact ctttcgagcc gctccaaatcg 5520
 gaggaggatg agagggaaatg atccgttccg gggggatcc tggggggatc caggaaattc 5580
 cctcgacgca tggccatatcg ggcacccggc gattacaacc ctccactgtt agagtccctgg 5640
 aaggacccttgg actacgtcccaatcgatgtt cttccatcgatc cttccatcgatc tggccaggcc 5700
 30 cctccatcgatc caccctccacg gaggaggagg acgggttgc tttccatcgatc tttccatcgatc 5760
 ttcgttccggc cggagctcgca caccacatcgatc tttccatcgatc tttccatcgatc 5820
 agccggacccgg caacggccctc tttccatcgatc caccacatcgatc tttccatcgatc 5880
 gtttggatgtt acttcgttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 5940
 gacggggatc tttccatcgatc aaggaggatc tttccatcgatc tttccatcgatc tttccatcgatc 6000
 tccttccatcgatc gacccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6060
 gtcaatgcac tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6120
 35 cgccggccaa ggcctccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6180
 cactaccggg acgtgtccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6240
 ctatccgtgg aaggaggctt taatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6300
 tatggggccaa aggacgttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6360
 tggaaaggact tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6420
 gagggtttttt tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6480
 40 ccagatttgg ggggttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6540
 cctcaggcccg tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6600
 ttccctggatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6660
 ttttttggatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6720
 ttttttggatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6780
 gggggccccc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6840
 ggtgtacttgc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6900
 tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 6960
 atctgtggaaa ggcggggacccggccaaatcgatc tttccatcgatc tttccatcgatc 7020
 atgactatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7080
 ataacatcatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7140
 tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7200
 actccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7260
 atgatccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7320
 ctagatgttgc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7380
 attcaacatcgatc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7440
 aatagggttgc tttccatcgatc tttccatcgatc tttccatcgatc tttccatcgatc 7500

EP 1 043 399 A2

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctt tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
5 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccgacc ccgctggtc atgtggtgcct tactcctact ttctgttaggg 7740
gtagggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttccctt ttttttttcc ttttttttcc tttggtgct ccatcttagc 7920
cctagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
10 tggectctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 7
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

<400> 7
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 ttctcacgca gaaagcgct agccatgcg ttatgtatgg tgcgtgcag cttccaggac 120
 cccccctccc gggagagcca tagtggctc cgaaaccggg gtagtacaccg gaattgcag 180
 gacgaccggg tccttcttc gatcaaccg ctaatgcgtt ggatgtttgg gctgcccc 240
 gcgagactgc tagccgatgt gtgtgggtc gcggaaaggcc ttgtggtaact gctgataagg 300
 gtgcttcgca gtccccggg aggtctcgta gacgtgcac catgagcagc aatcctaacc 360
 ctcaaaagaaa aaccaaaacgt aacaccacg ggcgcgccc gattgaaccaa gatggatgc 420
 acgcgggtc tccggccgtc tgggtggaga ggctattcgg ctatgactgg gcaacaacaga 480
 caatcggtc ctctgtatcc ggcgtgtcc ggctgtcgc gggggcgcg cgggttcc 540
 ttgtcaagac cgaccgttcc ggtgcctgaa atgaaactgcg gacgaggca ggcggcttat 600
 cgtggctggc cacgacggc gttccttgcg cagctgtcgtt cgacgttgct actgaagcgg 660
 gaagggactg gctgtatcc atcatgtgtc atgaaatgcg ggggtgtcat acgcttgatc 720
 ctctgtccg gaaatgtatcc atcatgtgtc atgaaatgcg ggggtgtcat acgcttgatc 780
 cggctacgtc cecatcgac cacaacgca aacatcgatc gacggcggc cgtactcgga 840
 tggaaacccggg tcttgcgtat caggatgtc tggacgaa gatcggggg ctgcgcgcg 900
 ccgaactgtt cgcaggctc aaggcggcga tgccgcggg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg ccgttttctt ggattcatcg 1020
 actgtggccg gtcgggtgtc gcgacccgt atcaggacat acgttgtgtt acccggtata 1080
 ttgtcttgcgaa gtttgggtc gaaatgggtc acgcttctt ctgttttac ggttgttcc 1140
 ctcccgattt cgcacgcattt gccttcttgc gccttcttgc cgatgttttc tgatgtttaa 1200
 cagaccacaa cgggttccctt ctgcgggtat caattccgc cctctccctt cccccccctt 1260
 aacgttactg gccgaagccg ctggaaataa ggcgggtgtt ctgtttgtcta tatgttatttt 1320
 tccaccatat tggcaatgtc aggccccggg aacctggccc tgcgtttcc 1380
 acgaggattt ctagggggtt ttcccttc gccaaagggg tgcacgggtt gttgaatgtc 1440
 gtgaaggaaag cagtcttgc ggaagcttgc tgaagacaaaa caacgtctgt acgcacccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgcctt gggccaaaaa gccacgtgt 1560
 taagatacac ctgcaaaaggc ggcacaaccc cagtgccacg ttgtgagttt gatagtgtt 1620
 gaaaagagtca aatggcttc ctcacggcgtt tcaacaagg ggctgaagga tgcccagaag 1680
 gtacccattt gtatgggtc tgatctggg ctcgggtca catgttttac atgtgtttag 1740
 tcgagggttaa aaaaacgttca gggccccccgga acacgggggg ctttgggggg 1800
 cacgataata ccatggggac gaaatctaaa cctcaaaaggaa aacccaaacg taacaccaacc 1860
 cggcccccac aggacgtcaa gttccggggc ggtgggtcaga tgcgtgggtt agtttactg 1920
 ttgcggcgca gggggcccaag gttgggtgtc cgccgcacta ggaagacttc cgagcgggtc 1980
 caaacctgtg gaaggccaca acctatcccc aaggctcgcc agccccgggg tagggccctgg 2040
 gtcagcccg ggtacccctt gccccttat gcaatgttgc gtttgggggtt ggcaggatgg 2100
 ctctgttcac cccgttgcgtc tcggccatgt tggggccca gggccccccg gctgttgc 2160
 cgcattttgg gtaagggtat cgataaccctc acgtgcgggt tgcgttgcattt catgggttac 2220
 attcggctcg tggcgccccctt ctttggggc gctggccaggg ccctggcgca tggcgtccgg 2280
 gtcttggagg acggcggtaa ctatgcaca gggaaatctgc cccgttgcctt ctgttttac 2340
 ttctttttgg ctgttgcgtc ctgttttgc atcccgatc cccgttgcattt gatgtcgcaac 2400
 gtatccggag ttttgcgttgc tcaatatacg gacggatgtt gatgttgc 2460
 gccggacatga tcatgtatcc cccgggtgc gtggccgtc tccggggggaa caactccccc 2520
 cgctgttggg tagcgttcac tcccacgtc gggccaggaa acgctacgtt cccactacg 2580
 acgatacgac gccatgtcga ttgtctgtt gggccgggtt ctctctgttc cgctatgtac 2640
 gtggggatc tctcgccgtc tggtttccctt gtcggccacg tggtttccatcc ctgcgttcgc 2700
 cggcacaacgca cagtcacggg ctgcgttgc tcaatatacg cggccacgtt gacagggttac 2760
 cgtatggctt gggatgtat gatgtactgg tcaatgttgc acgttccctt gatgtcgac 2820
 ttactccggat tcccacaaggc tgctgttgcgtt atgttggggg gggccattt gggaggctca 2880
 gcccccttgc cttactattt catggggggg aactgggtca aggttctgtat tgcgttgc 2940
 ctctttggcc gctgttgcgtt gggaaatctt gtcacggggg gacgtatggc caaaaacacc 3000
 ctccggatcc cgttccctttt ttcacccggg tcaatcccgatc tgtaaaacacc 3060
 aacggcgatc ggacatcaatc caggactgg ctgaatgtca atgacttccctt caaactctgg 3120
 ttcttgcgtc cgtgtttca cgtgcacaaatc tcaatgttgc tccggatgtcc agacggcgatc 3180
 gccagctgca gccccatcgatc cgcgttgcgtt cagggggtggg gccccatcac ttacaatgt 3240
 tcacacagct cggaccaggag gccttattgt tggcactacg cccccggcc gtcgtgttgc 3300
 gtaccggccgg cgcgggtgtt tggtccatgt tactgttca ccccaacggcc tgcgtgttgc 3360
 gggacgaccgg accgggttccg cgttccatcg tacatgttggg gggagaatgt gacggacgtg 3420
 ctgtttccatc acaacacggc ggcgcgcacaa ggcacactgtt ttgcgttgcac atggatgtat 3480
 agcaactgggt tcaccaaggac gtcggggggc cccccgtgttca acatcgggggg gatcggttac 3540

	aaaaccccttga	cctgccccac	ggactgtttc	cgaaagcacc	ccgaggccac	ttacaccaag	3600
5	tgtggttcg	ggcccttggtt	gacaccaga	tgcttgc	actaccata	caggcttgg	3660
	caactaccct	gactgtcaa	cttattacatc	ttcaaggta	ggatgtacgt	ggggggagt	3720
	gagcacaggc	tcaagccgc	atgaaattgg	actcgaggag	agegttgtaa	cctggaggac	3780
	agggacagat	cagagcttag	cccgctgctg	ctgtctacaa	cgaggatggca	gttattgccc	3840
	tgttccttca	ccaccctacc	ggctctgtcc	actggttga	tccatctcca	tcagaacgtc	3900
10	gtggacgtac	aatacctgt	cggtaggttgc	tctcttttgc	aatcaaattgg	3960	
	gagtagtgtc	tgttgc	ccttcttgc	gcccgcgc	gcgtctgtc	ctgcttgg	4020
	atgatgtgc	tgtatgtca	agctgaggcc	gccttagaga	acctgttggt	cctcaacgcg	4080
	gcatccgtgg	ccggggcgca	tggcatttctc	tccttctcg	tgttcttctg	tgctgcctgg	4140
	tacatcaagg	gcaggctgtt	ccctggggcg	gcataatggcc	tctacggcgt	atggccgcta	4200
15	ctcctgtcc	tgtggcttgc	accaccacga	gcatacgcc	ttgaccggga	gatggcagca	4260
	tctgtccggag	ggcggttgc	cgttagtctg	atacttgc	ccttgc	gcactataag	4320
	ctgttcctcg	ctaggctcat	atggtgggtt	caatattttt	tcaccaggggc	cgagggcacac	4380
	ttgcaagtgt	ggatcccccc	cctcaacgtt	cgggggggcc	gcatgtccgt	catccttc	4440
	acgtgcgcga	tccacccaga	gctaattttt	accatcacca	aaatcttgc	cgccataactc	4500
20	ggtccactca	tgggtctca	ggctggtata	accaaattgc	cgtacttgc	gcgcgcacac	4560
	gggcttcatc	gtgcgtcat	gctggcg	aaagggtctg	gggggtcatta	tgtccaaatg	4620
	gctctcatga	agttggccgc	actgacagg	acgtacgtt	atgaccatct	cacccactg	4680
	cgggactggg	cccacgcggg	cctacagac	tttgcgttgc	cagttgagcc	cgtcgtctc	4740
	tctgatatgg	agaccaagg	tatcacttgc	ggggcagaca	ccggccgt	tggggacatc	4800
25	atcttggcc	tggcgctc	cgcccgagg	ggggaggaga	tatcatctgg	accggcagac	4860
	agccttgaag	ggcagggttg	gegacttctc	ggcccttata	ccggcttactc	ccaacagacg	4920
	cgaggcttac	ttggctcat	catcaactgc	ctcacaggcc	gggacaggaa	ccaggtcgag	4980
	ggggagggtcc	aagtggctc	caccgaaca	caattttcc	tggc	gacactg	5040
	gtgtgttgg	ctgtctatca	tggtgccggc	tcaagaccc	ttggccggccc	aaaggggcca	5100
30	atcacccaaa	tgtacaccaa	tgtggaccag	gacccgtc	gtctggca	gccccccggg	5160
	gcgcgttctt	tgcacccat	cacctggcc	agtcggacc	tttacttgg	cacgaggat	5220
	gccgatgtca	ttccggtgc	cegggggggc	gacagcagg	ggagccact	ctccccca	5280
	cccggtctct	acttgaagg	cttccggc	ggtccactgc	tctggccctc	ggggcatgt	5340
35	gtgggcatct	ttgggtctgc	cgtgtgcacc	cgagggttg	cgaggccgt	ggactttgt	5400
	cccggtcagt	ctatggaaac	catatcg	ttcccggt	tcggacaa	ctcgcccc	5460
	ccggccgtac	cgacgacatt	ccagtgcc	catctacacg	ccctactgg	tagcgc	5520
	agcaactaagg	tgcggctgc	gtatgc	caagggtata	agggtcttgc	cctgaacccg	5580
	tccgtcgccg	ccacccatgg	tttccggc	tatatgtct	aggcacatgg	tatcgacc	5640
40	aacatcagaa	ccgggttgg	gaccatcacc	acgggtgccc	ccatcacgt	ctccacat	5700
	ggcaagttc	ttggcgac	tgggtgtct	ggggggccct	atgacatcat	aatatgtat	5760
	gagtgcact	caactgactc	gaccatc	ctggc	atcgactt	ggaccaagcg	5820
	gagacggctg	gagcgcact	cgtcgtctc	gccacccgt	cgctccggg	atcggtcacc	5880
	gtgcacacatc	caaacatcga	ggagggtgt	ctgc	ccatggaaat	cccttttat	5940
45	ggcaaaggca	tecccatcgc	gaccatca	ggggggaggc	acccat	ctgcatttc	6000
	aagaagaaat	gtgatgacgt	cgccgc	aaag	ctatccgg	tcgactaa	6060
	tattaccggg	gccttgatgt	atccgtcata	ccaact	agcgt	tgtctgt	6120
	acggacgctc	taatgacggg	cttaccggc	gattc	act	gtcaata	6180
	tgtgtcaccc	agacagt	cttcacgt	gaccggac	tc	accat	6240
	gtgcaccaag	acgcgtgtc	acgtc	cgccgagg	ggact	ttgg	6300
50	ggcattttaca	gttttgc	tccaggagaa	cgccctcg	gtat	ttccctcg	6360
	ctgtcgagat	gtctatgacgc	gggtgt	tggta	ccac	cgagac	6420
	gttaggttgc	gggttta	aaacacacca	gggt	ccat	ctgg	6480
	ttctgggaga	ggcttta	accat	cgac	ccat	ttgg	6540
	aggcaggcag	gagaca	accat	ctcg	atctt	ccat	6600
	gtcaggcgtc	cacccatc	gtgg	ggac	atgt	ggc	6660
	acgctgcacg	ggccaac	cct	gtgt	ggc	ggact	6720
	accacacacc	ccataacca	at	acat	ccgt	ggag	6780
	acgacgaccc	gggtgtgt	aggc	ggac	tc	ggcc	6840
	acaggcagcg	tggt	cattt	ggc	at	ccat	6900
	gacagggaag	tcctt	accat	gg	ttt	ccat	6960
	tacatcgaa	aggaaat	gtc	gtt	ccat	ccat	7020
	caaacagcca	ccaac	aaat	ccat	ccat	ccat	7080
	ctcgaagcct	tctggc	ggat	ttt	ccat	ccat	7140
	ggcttgc	cttgc	ccat	gg	ttt	ccat	7200
	attaccagcc	cgct	ccat	ccat	ccat	ccat	7260
	ggcgcccaac	ttgt	ccat	ccat	ccat	ccat	7320
	gogggttgc	gca	ccat	ccat	ccat	ccat	7380
	gcaggggtgg	ca	ccat	ccat	ccat	ccat	7440
	gaggacacttgc	tta	ccat	ccat	ccat	ccat	7500

	gtgtgcgcag	cgataactgcg	tccgcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
	aaccggctga	tagcgttcgc	ttcgcggggt	aaccacgtct	cccccacgca	ctatgtgtct	7620
5	gagagcgcacg	ctgcagcagc	tgtcaactcg	acctctctcta	gtcttaccat	cactcagctg	7680
	ctgaagaggc	ttcaccatgt	gatcaacgag	gactgtccca	cgccatgtctc	cggctcgtgg	7740
	ctaagagatg	tttgggattt	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaaagctcc	tgccgcgatt	gccgggagtc	cocttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaacgggtc	catgaggatc	gtggggctta	ggacactgtag	taacacgtgg	7980
10	catggaacat	tccccattaa	cgcgtacacc	acggggccct	gcacgcctc	ccggcgcca	8040
	aattattctt	gggcgcgtgt	gcgggtggct	gtcgaggagt	acgtggaggt	tacgcggggt	8100
	ggggattttcc	actacgtgac	gggcgtgacc	actgacaacg	taaagtgcctt	gtgtcagggtt	8160
	ccggcccccgg	aatttttcac	agaatgtggat	gggggtcggt	tgcacaggtt	cgctccagcg	8220
	tgc当地	tcctacgggg	ggggatcaca	ttcctgggtc	ggctcaatca	ataccctgggt	8280
	gggtcacagc	tccccatgcgg	ggccgaaaccg	gacgttagcag	tgctcacttc	catgtctcacc	8340
15	gacccttccc	acattacggc	ggagacggct	aacgttaggc	tggccagggg	atctccccc	8400
	tccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	acccgtcatg	actccccgg	cgctgaccc	atcgaggcc	acctccctgt	gcggcaggag	8520
	atggccggga	acatcaccc	cgtggagtca	aaaaataagg	tagtaatttt	ggactctttc	8580
	gagccgctcc	aagcggagga	ggatgagagg	gaagtatccg	tcccgccgga	gatccctgcgg	8640
	aggccagga	aatttttc	agcgtgccc	atatgggcac	gcccggatta	caaccctcca	8700
20	ctgttagagt	cctggaaagg	cccgactac	gtccctccag	tgttacacgg	gtgtccattt	8760
	ccgcctgcca	aggcccctcc	gataccac	ccacggaggg	agaggacgtt	tgtccctgtca	8820
	gaatctaccg	tgttttctgc	cttggcgag	ctcgccacaa	agaccttcgg	cagctccgaa	8880
	tcgtcgccg	tcgacagcgg	cacggcaac	gccttcctg	accagccctc	cgacgcggc	8940
	gacgcggat	ccgacgttga	gtcgtaactc	tccatgcccc	cccttgaggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggct	acgtaagcg	aggaggctag	tgaggacgtc	9060
25	gtctgtgtc	cgatgttca	cacatggaca	ggcgccctga	teacgcctat	cgctgcggag	9120
	gaaacaaggc	tgccgtca	tgcactgagc	aacttttgc	tccgttacca	caacttggtc	9180
	tatgtacaa	catctcgca	cgcaagcctg	cggcagaaga	aggtcacctt	tgacagactg	9240
	caggctctgg	acgaccacta	ccgggacgt	ctcaaggaga	tgaaggcgaa	ggcgccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtta	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaggac	gtccggacc	tatccagcaa	ggccgttaac	9420
30	cacatccgct	ccgtgtggaa	ggacttgtc	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggcaa	aaaatggat	tttctcgct	caaccagaga	agggggcccg	caagccagct	9540
	cgcttatacg	tattttccaga	tttgggggtt	cgtgtgtcg	agaaaaatggc	cctttacat	9600
	gtggcttcca	ccctccctca	ggccgtgtat	ggcttctat	acggattca	atactctct	9660
	ggacagcggg	tcgagttct	gtgtatgcc	tgaaaagcga	agaaatgccc	tatgggcttc	9720
35	gcatatgaca	cccgctgtt	tgactcaac	gtcaactgaga	atgacatccg	tgttgaggag	9780
	tcaatctacc	aatgttgc	cttggccccc	gaagccagac	aggccataag	gtcgctcaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cggtcgcgcg	cgagcggt	actgacgacc	agctgcggta	ataccctcac	atgttatttt	9960
	aaggccgtcg	cggccgtgt	actccggact	tgccatggat	gcaatgtat	cgtatgcgg	10020
	gacgacccctg	tcgttatctg	tggaaagcgt	ggggacccaa	aggacgggg	gaggctacgg	10080
	gccttcacgg	aggctatgac	tagatactc	gccccccctg	gggacccccc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgtcc	tccaatgtgt	cagtcgcgca	cgatgcacat	10200
	ggcaaaaggg	tgtactatct	caccgtgac	cccaccaccc	cccttgcgcg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgccc	10320
	accttgcggg	caaggatgt	cctgtatgact	cattttttct	ccatcccttc	agctcaggaa	10380
	caacttgcgg	aagccctaga	ttgtcagatc	tccggggctt	gttacttccat	tgagccactt	10440
	gacctaccc	agatcattca	acgactccat	ggccttagcg	cattttccat	ccatagtatc	10500
45	tctccagggt	agatcaatag	ggtgggttca	tgcctcaggaa	aacttgggg	accggccctt	10560
	cgagtctgga	gacatcgccc	cagaagtgtc	cgcgcctaggc	tactgtccca	ggggggggagg	10620
	gctgcceactt	gtggcaagta	cctcttcaac	tgggcagtaa	ggaccaagct	caaactcaet	10680
	ccaatccccc	ctgcgtccca	gttggattt	tccagctgg	tcgttgcgtt	ttacagcggg	10740
	ggagacatata	atcacagcct	gtctcgtcc	cgaccccgct	ggttcatgtt	gtgcctactc	10800
50	ctactttctg	taggggtagg	catctatcta	ctccccaacc	gatgaacggg	gagctaaaca	10860
	ctccaggccca	ataggeccatc	ctgttttttt	cccttttttt	tttttttttt	tttttttttt	10920
	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	10980
	tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gtttgactgc	11040
	agagagtgt	gatactggcc	tctctgcaga	tcaagt			11076

<210> 8
 <211> 8001
 <212> DNA
 <213> Hepatitis C Virus

5 <400> 8
 gcacggccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
 ttttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cttccaggac 120
 ccccccctccc gggagagcca tagtggctcg cgaaaccggg gatgacaccg gaattggcag 180
 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttg gcgtgcccc 240
 gcgagactgc tagccgagta gtgttgggtc gceaaaggcc ttgtgttact gcctgatagg 300
 gtgcgtcgaa gtggcccccggg aggtctcgta gaccgtgcac catgagcacaa aatccctaaac 360
 ctcggaaa aacccaaacg aacccaaacg ggogcgccat gatgaaacaa gatggattgc 420
 acgcagggttc tccggccgtc tgggtggaga ggctattcgat ctatgactgg gcacaacaga 480
 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttctt 540
 ttgtcaagac cgacctgtcc ggtgcctga atgaactgca ggacgaggca ggcggctat 600
 cgtggctggc cacgacggc gttcctcgca cagctgtctc cgacgttgc actgaacgg 660
 gaagggactg gtcgttattt ggcgaagtgc cggggcaggat ttcctgtca ttcacctt 720
 ctccctgcga gaaaggatacc atcatgctg atgaatgcg gggctgtcat acgcttgcata 780
 cggttacctg cccatcgac caccaacgca aacatcgcat cgacgagca cgtactcgga 840
 tggaaaggccg ttttgcgtat caggatgatc tggacgaaaga gcatcagggg ctcgcgcag 900
 ccgaactgtt cccagggttc aaggcgcgc tggccgcgg cgaggatctc gtcgtgaccc 960
 20 atggcgatgc ctgcgtcgca aatcatatgg tggaaaatgg cgcttttttggattcatcg 1020
 actgtggccg gctgggtgtc gccggacgcgt atcaggacat agcgttgcgt acccgatata 1080
 ttgtcaaga gcttggccgc gaatgggtc accgttctt cgtctttac ggtatcgccg 1140
 ctcccgattt gcacgcgcata gccttcttate gccttcttgc ctagtttttc tgagttaaa 1200
 cagaccacaa cgggtttccctt ctageggat ctaatccgc cctcttcctc cccccccctt 1260
 aacgttactg gccgaagcgc ctggaaataaa gggcgtgtc tggatgttata tatgttattt 1320
 tccaccatata tccgcgtctt tggcaatgtt agggccggaa aacatggccc tgccttcgtt 1380
 25 acgagcattc ctagggttctt tccccctctc gccaaggaa tgcacggatc gttgaatgtc 1440
 gtgaaggaaag cagttccctt ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 tgcaggcagc ggaacccccc acctggcgac aggtgcctct gccggccaaaa gccaatgtca 1560
 taagatacac ctgcacaaaggc ggcacaaaccc cagtgccacg tggatgttgc gatagttgt 1620
 gaaagagtca aatggctctc ctcacgcgtat tcaacaaagg ggctgtggaa tgcccaaaag 1680
 gtacccattt gtatggatc tgatctggg ctcgggtgtca catgtttac atgtttag 1740
 tcgaggtaaa aaaacgtcta gggcccccga accacggggc cgtgttttc ctttggaaaa 1800
 cacgataata ccatggcgcc tattacggc tactccaaac agacggcgagg cctactgtgc 1860
 30 tgcatcatca ctgcctcag accggggac aggaaccagg tggatggggg ggtcaatgt 1920
 gtcctccaccg caacacaaatc ttctctggc acctgcgtca atggatgttgc ttggactgtc 1980
 ttcatgtgtt cggcgtcaaa gaccctgtcc ggcggaaagg gccaatcactc ccaaatgtac 2040
 accaatgtgg accaggaccc cgtgggttcc caagegecccc cggggggcgg ttccttgaca 2100
 ccatgcacctt gggcgacgttcc gggccgttcc ttggatgttgc ggcacggccca tgcatttcgg 2160
 35 gtgcgcggc gggggcggcagc cggggggcggc ctactctcc cagggccctt ctcctacttg 2220
 aagggtctt cggccgttcc acgtctgtc ctcctggggc acgtgtggg catcttccgg 2280
 gtcgcgtgtt gcaacccgggg ggttgcgttcc ggggtggact ttgttccgtt cggatgtat 2340
 gaaaccacta tgcggtcccc ggttccatcg gacaactcgat cccctccggc cgtaccgcag 2400
 acattccagg tggcccatat acacgccttact gtttgcgttgc gcaagggcact taagggtggc 2460
 gtcgcgtatg cagggccaaagg gtatgggtt cttgttccatcg acccggttgcgtt cggcccccacc 2520
 40 cttaggttcg gggcgatata tgcgttccatcg catgttgcgttgc acccttaatcat cagaatcggtt 2580
 gtaaggacca tcaaccacggg tggcccccattt acgttacttca cctatggca gtttcttgac 2640
 gacgggtt gtcgtggggg cgcctatgtac atcataatat gtatgttgcgtt ccactcaact 2700
 gactcgacca ctatcttggg catcgccaca gtcctggacc aagggggagac ggctggagcg 2760
 cgaactgtcg tgcgtccac cgcgttgcgttcc cggggatcgatccatcgatgttgc acatccaaac 2820
 atcgaggagg tggctgttcg cagactggaa gaaatccctt tttatggca gcccatcccc 2880
 atcgagacca tcaagggggg gaggccatcc atttttgcctt atccaaagaa gaaatgtat 2940
 45 gagctcgccg cgaagctgtc cggccctcgaa ctcaatgtcg tagcatatata cggggccctt 3000
 gatgtatccg tcatatccaaatc tagcggtggac gtcattgtcg tagcaacggc cgctctaaatg 3060
 acgggttta cgggtgtactt cgcactgttcc atcgactgtca atatcatgttgc cacccttggaca 3120
 gtcgacttca gtcgtggacc gacccatccacc attggatgtca cggccgttgc acacggccgg 3180
 gttgtacgtt cgcacggccgg agggcggactt ggttggggca ggtatggggcat ttacagttt 3240
 gtgactccag gagaacggcc ctcggccatcg ttcgttccctt cgggttgcgtt cgagtgctat 3300
 50 gacgcgggtt gtgcgttgcgtt cgcacggccatcg cccggccgaga cctcgtttag gttgcgggtt 3360
 tacctaaaca caccagggtt gcccgttgc caggaccatcg tggatgttgcgtt ggagagcgctc 3420
 ttacaggcc tcaaccacat agacgccttccat ttcgttccctt agactaaagca ggcaggagac 3480
 aactccctt acctggatcg acgttgcgttgc cggggccatcg ggctccactt 3540

5 ccatcggtgg accaaatgtg gaagtgttcc atacggctaa agcctacgt gcacgggcca 3600
 acggcccgtgc tgtataggct gggagccgtt caaaacgagg ttactaccac acaccccata 3660
 accaaataca tcatggcatg catgtcggt gacctggagg tcgtcaacgag cacctgggtg 3720
 ctggtaggcg gagtccttagc agctctggcc gcgtattggc tgacaacagg cagcgtggtc 3780
 attgtggca ggatcatctt gtccggaaag cggccatca ttcccgcacag ggaagtctt 3840
 taccgggagt tcgatagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
 atgcagctcg cggaaacaat caaacaagg gcaatccggg tgctgeaaac agccaccaag 3960
 caagcggagg ctgctgtcc ctgggtggaa tccaatggc ggaccatcg agccttctgg 4020
 gcgaagcata tggaaattt catcagcggg atacaatatt tagcaggctt gtccactctg 4080
 10 cctggcaacc cccgcataatc atcaactatgc gcattcacag cctctatcac cagcccgctc 4140
 accacccaaat atacccttc gtttaacatc ctggggggat gggtgccgc ccaacttgc 4200
 cctccacgc ctgttttcg tttcttaggc gccggcatcg ctggagccgc tggtgcagc 4260
 ataggcttc ggaagggtgc ttggatattt ttggcagggtt atggagcagg ggtggcaggc 4320
 gcgcctgtgg cttttaaggt catgagcggc gagatgcctt ccacccggaa cctggtaac 4380
 ctactccctg ctatccctc ccctggccgc ctatgtcg gggtcgtgt cgccagcgata 4440
 ctgcgtcgcc acgtggggcc aggaggaggd gctgtcgactt ggtgaaaccg gctgatagcg 4500
 ttgcgttcg gggtaacca ctgttcccc accgcactatg tgccttagag cgacgtcgca 4560
 gcacgtgtca ctcaatctt ctctatgtt accatactc agctgtcgaa gaggcttcac 4620
 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680
 gattggatat gcacgtgtt gactgtttt aagacatcgcc tccagtcacaa gtcctgtccg 4740
 cgattgggg gagtccctt ctctcatgt caacgtgggtt acaaggaggat gttggggggc 4800
 gacggcatca tgcaaaacc cttggccatgtt gggccacaga tcacccggaca tggaaaaaac 4860
 ggttccatga ggatcggtt gccttaggacc ttgtatataca cgtggcatgg aacattcccc 4920
 20 attaacgcgt acaccacggg ccctgcacgc ccctccccgg cgccaaatata ttctagggcg 4980
 ctgtggccgg tggctgtca ggatgtcg gagggtacgc ggggtggggg ttccactac 5040
 gtgacggcca tgaccatcg cgaatggaa tgccctgtc aggttccggc ccccaatttc 5100
 ttcacagaag tggatggggt ggggttgc acgtacgtcc cagcgtcaaa accccctccct 5160
 cgggaggagg tcacatctt ggtcggtc aatcaatacc ttgttgggtc acagctcccc 5220
 tgcgagcccg aaccggatgt acaactgtcc acttccatgc tcacccgaccc ctccacatt 5280
 acggccggaa cggctaaagg taggtcgccg aggaggatctc ctccccccctt gcccagctca 5340
 tcagcttagc agctgttcg gccttcctt aaggcaacat gcaatcccg tcatgactcc 5400
 ccggacgtcg acctcatcg ggcacccctc ctgtggccgc aggagatggg cgggaacatc 5460
 acccgctgg agtcagaaaa taaggtaga attttggact ctttcgagcc gtcacccatcg 5520
 gaggaggatg agagggaaatg atccgttccg ggggagatcc tgccggaggtc cagggaaattc 5580
 cctcgacgca tgcccatatg ggcaccccg gattacaacc ctccactgtt agactctgg 5640
 aaggaccctgg actatccctt ccgtatggta cccgggtgtc cattccggcc tgccaaaggcc 5700
 30 cctccgatcc caccttcacg gggaaaggagg acgggtgtcc tgcacaaatc taccgtgtct 5760
 tctgccttgg cggagctcg cacagagacc ttccggcagct ccgaatcgcc ggcgtcgac 5820
 agccgcacgg caacggccct tcctgaccag ccctccgacg acggcgcacgc gggatccgac 5880
 gttgagtcgt actccctcat gccccccctt gggggggggc cgggggatcc cgatctcage 5940
 gacgggttcc ggtctatcg aaggcggaggat gctatgttggggc acgtcgctgt ctgcgtatg 6000
 tcctacacat ggacaggccg cctgtatcacc ccatgcctg cggaggaaac caagctgccc 6060
 atcaatgcac tgagcaactc ttgtccctt caccacaact ttgtctatgc tacaacatct 6120
 cgcacgcacaa gcctggccaa gaagaagggtt acctttgaca gactgcaggat cctggacgac 6180
 cactacccgg acgtgtccaa ggagatggaa gcaaggccgtt ccacatgtt ggctaaactt 6240
 ctatccgtgg aggaaggctt taatgtcgac cccccacatt cggccatgtc taaatggc 6300
 tatggggccaa aggacgtccg gaaacctatcc agcaaggccg ttaaccatcat cccgtccgt 6360
 tggaaaggact tgctggaaaga cactggagaca ccaattgaca ccacccatcat ggcaaaaaat 6420
 gaggttttt gcgtccaaagg agagaagggg gggccgacac cagctccgtt tattgttattc 6480
 40 ccagatttgg ggggttgcgtt gtggcgagaaa atggccctttt acgtatgttgc ctccacccttc 6540
 cctcaggcccg tggatgggttccatcggatc ttccatcggatc tccctggaca ggggtcgag 6600
 ttccctggta atgcctggaa afgaaagaaa tgccctatgg gcttcgcata tgacacccgc 6660
 tggggggactt caacgttcac tgagaatgac atccgtgttgg aggaggtaat ctaccaatgt 6720
 tggactgtgg cccccccatcg cagacaggcc ataagggtcc tcacagatcg gctttacate 6780
 gggggccccc tgactaattt taaaggggcacttgcgggtt atccgggtg cccgcgcgac 6840
 ggtgtactcg cggacccatcg cggatataatc ttccatcgatgtt ggtggggac cgtcgccggcc 6900
 tggactgtgg cggactgtccaa ggactgcacg atgtctcgat gggggacatc ctttgtcggt 6960
 atctgtggaa ggcggggac ccaagaggac gaggccggcc tacgggcctt cagggaggat 7020
 atgactgtccaa ccctggggac cggcccaaaatc cagaatcgaa cttggatgtt 7080
 ataacatcat gtcctccaa tggatcgatc ggcacatcgatc tggatggcaaa aagggtgtac 7140
 tatttcaccc tggaccccaac cacccttgcgtt gggggggatc cgtggggatc agctagacac 7200
 actccactca attcctggctt aggaacatc atcatgtatc cggccacccctt gttggcaagg 7260
 50 atgatctgtca tgactcattt ctctccatc ttcttagtgc aggaacaact tgaaaaagcc 7320
 ctagatgttcc agatctacgg ggcctgttac tccatggac cacttgaccc acctcagatc 7380
 attcaacacatc tccatggccct tcaatccatc gttactctcc aggtggatgc 7440
 aatagggtgg ctcatcgatc cggatggatc ggggttccatc ctttgtcgatc ctggagacat 7500

EP 1 043 399 A2

5 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctt tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
10 agcctgtctc gtgcccgacc ccgctgggtc atgtggtgcc tactcctact ttctgttaggg 7740
gtaggcatct atctaactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt ttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttcctt ctttttttcc ttttcttcc tttggtggtt ccatcttagc 7920
cctagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
tggcctctt gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 9
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 5
 <400> 9
 gccagcccc gattgggggc gacactcac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgctc agccatggcg ttagtatag tgcgtgcag cttccaggac 120
 ccccccctccc gggagagcca tagtggctg cggAACCGGT gagtacaccg gaatttgcag 180
 gacgaccggg tccttcttg gataaaccgg ctcaatgcct ggagatttgg gcgtgcggcc 240
 10 gcgagactgc tagccgagta gtgtgggtc gcggAAAGCC ttgtgtact gcctgatagg 300
 gtgttgcga gtggccggg aggtctcgta gaccgtgcac catgagcacg aatccctaaac 360
 ctcaaaagaaa aacccaaacg aacccaacg ggcgcgcct gattgaacaa gatggattgc 420
 acgcagggtc tccggccgt tgggtggaga ggctattcg tcatgactgg gcacaacaga 480
 caatcggtc ctctgtatgcc gccgttcc ggctgtcgc gcaggggcgc cccgttctt 540
 ttgtcaagac cgaccgttcc ggtgcctga atgaactgcg ggacgaggca gcgcggctat 600
 15 cggtgcgtgc cacgacgggc gttcttgcg cagctgtgtc cgacgtgtc actgaagcgg 660
 gaaggggactg gctgttattc ggcgaagtgc cggggcaggaa ttcctgtca ttcacccctt 720
 ctctgcgcga gaaatgtatcc atcatggcg atgcaatgcg gggctgcatt acgcttgatc 780
 cggttacccgt cccatcgac caccacggca aacatcgatc cgagcgagca cgtactcgga 840
 tggaaagccgg tcttgcgtat caggatgtc tggacgaaaga gcatcagggg ctcgcgcag 900
 20 ccggactgtt cggccggctc aaggccggca tgcccgcgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgg aataatcg tggaaaatgg ggttttttgc ggtttcatcg 1020
 actgtggccg gctgggtgtg gccggccgt atcaggatc acgcttgatc acccggtata 1080
 ttgtcaaga gcttggccgc gaatgggtg accgcttctt cgtgttttac ggtatcgccg 1140
 ctcccgatc gcagcgatc gccttctatc gccttcttgc ctagtttttc ttagttttaa 1200
 25 cagaccacaa cgggtttccct ctgcggat caattccgc ccttccttcc cccccccctt 1260
 aacgttactc gccgaaggcc cttggataaa tgggggtgtg cttttgtata tatgttattt 1320
 tccaccatc tgccgttcc tggcaatgtg aggggccggaa aacctggccc tgcgttctt 1380
 acgagcatc ctagggttct ttcccctctc gccaaaggaa tgcaggatgc gttgaatgtc 1440
 gtgaaggaag cagtttcttgc ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
 30 tgcaggcagc ggaacccccc acctggcgac aggtgccttgc gggccaaaa gccacgtgtt 1560
 taagatacac ctgcaaggc ggcacatccc ctaggcacg tgggtggat gatagttgt 1620
 gaaagagtc aatggcttc ctcacacgtt ttcacaacgg ggctgaaggaa tgcccaagaag 1680
 gtacccatt gtatggatc tgatctgggg ctcgggtcgc catgttttac atgtgtttag 1740
 tgcagggtttaaaa acgtcttca ggcggccggaa accacggggaa cttttttttc ctttggaaaa 1800
 35 caccataa ccatggccac gaatctaaa ctcacacggaa aaacccaaacg taacaccaac 1860
 cggccggccac aggacgtcaa gttccggggc gttgggtcaga tcgtcggtgg agtttacctg 1920
 ttgcggcgcga gggggccggc gttgggtgtg cggcgcacta ggaagacttc cgagcggtcg 1980
 caacctcgta gaaggcgacaa acctatcccc aaggctcgcc agcccgagggg tagggctgg 2040
 gctcagccgg ggttcccttgc gcccctctat gcaatgggg gcttgggtgt ggcaggatgg 2100
 ctccgttac cccgtgttgc tcggcctagt tggggccccc cggacccccc gctgttgcgt 2160
 40 cgcacattttgg gtaagggtatc ctaggttcc acgtgcgggt tcgttgcattt catgggttac 2220
 atcccgctcg tcggccccc ctttgcggcgc gtcgttgcggg ccttgcgcga tggcgccgg 2280
 gttctggagg acggcggttca ctagtgcaca gggaaatctgc cccgttgcgc tttttctatc 2340
 ttccctttgg ctttgcgttgc ctgtttgcacc atcccagttt cccgttatgc agtgcgcac 2400
 gtagtccggag tgcgttgcgttgc ctagtgcaca gggaaatctgc cccgttgcgc tttttctatc 2460
 45 gccggatgtca tcatgtatcc cccggccggc gttggggatgg ttcggggatggaa caactctcc 2520
 cgcgtgtgg tagcgttcc tcccaacgttgc gggccggaa acgttgcgttgc cccactacg 2580
 acgatacgtac gccatgtcga ttgtgttgc gggccggcgttgcgttgc cccactacg 2640
 gtggggatgtc tgcgttgcgttgc tttttcccttgc tgcgttgcgttgc tttttcccttgc 2700
 cggccacggaa ctttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 2760
 cgtatgggttgc gggatgtatgt gataactgg tccatctacg cggccctgttgc ggtatcgac 2820
 ttactccggatgtc tcccaacgttgc gttgggtgttgc tttttcccttgc tttttcccttgc 2880
 gccggccggcgttgc ctttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 2940
 50 ctctttgcgttgc gggccggcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 3000
 ctccggatgtca tttttcccttgc tttttcccttgc tttttcccttgc tttttcccttgc 3060
 aacggccatccaa ctttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 3120
 ttcccttgcgttgc ctttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 3180
 gccggccggcgttgc tttttcccttgc tttttcccttgc tttttcccttgc tttttcccttgc 3240
 tcacacatgttgc tttttcccttgc tttttcccttgc tttttcccttgc tttttcccttgc 3300
 gtaccggccgg ctttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc 3360
 gggacggccgg accggccgttgc tttttcccttgc tttttcccttgc tttttcccttgc 3420
 ctgttgcgttgc tttttcccttgc tttttcccttgc tttttcccttgc tttttcccttgc 3480
 agcaactgggttgc tttttcccttgc tttttcccttgc tttttcccttgc tttttcccttgc 3540

	aaaacacctga	cctggcccac	ggactgcttc	cggaagcacc	ccgaggccac	ttacaccaag	3600
5	tgtggttcg	ggccttgggt	gacaccaga	tgttggtcc	actaccata	caggcttgg	3660
	caactaccct	gcaactgtcaa	cttaccatc	ttaagggtt	gatgtacgt	ggggggagt	3720
	gagcacacggc	tgaaggccgc	atgcaatttg	actcgaggag	agegttgtaa	cctggaggac	3780
	agggacagat	cagagcttag	cccgetctg	ctgtctacaa	cgagtgcc	gttattgccc	3840
	tgttccttca	ccaccctacc	ggctctgtc	actggttga	tccatctcca	tcagaacgtc	3900
	gtggacgtac	aatacctgt	cggtataagg	teggcggtt	tctccttgc	aatcaaatgg	3960
	gagtagtgtcc	tgttgcttt	cattcttctg	ggggacgcgc	gctgtgtgc	ctgcttgg	4020
10	atgatgctgc	tgatagctca	agctgaggcc	gccttagaga	acctgtgtt	cctcaacgc	4080
	gcateccgtgg	ccccggcga	tggcattctc	tccttccctg	tgttcttctg	tgctgcctgg	4140
	tacatcaagg	gcaggctgt	ccccgggcg	gcatatgcc	tctacggcg	atggccgc	4200
	ctctgtctcc	tgtggcggt	accaccaga	gcatacgc	ttgaccgg	gatggcagca	4260
	tcgtggggag	ggcggtttt	cgttaggtct	atacttttga	ccttgcacc	gcaataaag	4320
	ctgttctctcg	ctaggctcat	atggtgggtt	caatatttt	tcaccagg	cgaggcacac	4380
15	ttgcaagtgt	ggatcccccc	cctcaacgtt	ccccggggcc	gcatgtccgt	catccctc	4440
	acgtgcgcg	tccaccaca	gtaatacttt	accatcacca	aaatcttgc	cgccataactc	4500
	ggtccactca	tgtgtctca	ggctgttata	accaaaatgc	cgtaactcg	gcccgcacac	4560
	gggttcattc	gtgtcatgt	gtggcgccg	aagggtgtc	gggggtcatta	tgtccaaatg	4620
	gtctctatga	agttggccgc	actgacaggt	acgtacgtt	atgaccatct	cacccactg	4680
	cgggacttgg	ccccacgcgg	cctacgagac	cttgcgggtt	cagtgtgagc	cgtcgcttc	4740
	tctgtatgg	agaccaagg	tatcacttgc	ggggcagaca	ccgcgcgt	tggggacatc	4800
	atcttggcc	tccctgtctc	cgcccg	ggggggaga	tacatctgg	accggcagac	4860
20	agccttgaag	ggcagggggt	gcgactctc	gcgccttata	cgccctactc	ccaacagacg	4920
	cgaggcctac	ttggctgtc	catcaactagc	ctcacaggcc	gggacaggaa	ccaggtcgag	4980
	ggggagggtcc	aagtgttctc	cacccaaaca	caatcttcc	tggcgcac	cgtcaatg	5040
	gtgtgttgg	ctgtctata	tgtgtcg	tcaagacatc	ttggccggcc	aaaggggcc	5100
	atcacccaaa	tgtacaccaa	tgtggacc	gacccgtcg	gtggcaagc	gccccccgg	5160
	gcgcgttctc	tgacaccatg	cacctgcgg	agctcgacc	tttacttgg	cacgaggcat	5220
25	gccgatgtca	ttccggtgc	ccggggggc	gacagcagg	ggagcctact	ctccccca	5280
	cccgctctct	acttgaagg	cttctggcc	gttccactgc	tctggccctc	ggggcacg	5340
	gtgggcatct	ttccggctc	ctgtgtc	cgagggttgc	cgaggcggt	ggactttgt	5400
	cccgtcgagt	ctatggaaac	cactatcg	ccccggct	tccagg	ctcgcccc	5460
	ccggccgtac	cgccggat	ccagggtgg	catctacac	ccccctact	tagggcaag	5520
	agcactaagg	tgccggctc	gtatgcag	caagggtata	aggtgtt	cctgaacc	5580
	tccgtcgcc	ccacccatgg	tttccggcc	tatatgtcta	aggcaccat	tategacc	5640
30	aacatcaga	tccgggtaa	gatccatc	acgggtgccc	ccatcag	ctccac	5700
	ggcaagttc	ttggcgcac	ttgtgtct	gggggcct	atgacat	aatatgtat	5760
	gagtgcact	caactgact	gaccactat	ctggc	gcacgtt	ggaccaag	5820
	gagacggctg	gagcgcact	cgtcg	gcacccgt	cgccccc	atcggtc	5880
	gtgccacatc	caaacatcg	ggagggttgc	cttgc	ctgg	aaaaat	5940
	ggcaagacca	tcccccata	gatccat	ggggggag	acccat	ctggcatt	6000
	aagaagaaat	gtgtgtat	cgccgc	cttgcgg	tcggact	tgtgt	6060
35	tattaccggg	gccttgatgt	atccgtata	ccaaact	cgatgt	tgtgt	6120
	acggacgctc	taatgacgg	tttacccgt	gactcg	cagtgt	ctgaaata	6180
	tgtgtcaccc	agacatcg	tttccgc	gaccgc	taccat	gtgacg	6240
	gtgcacaa	acgcgtgtc	acgtcg	ccggagg	gatgttgc	ggcaggat	6300
	ggcatttaca	gtttgt	tccagg	ccgcct	gatgttgc	ttcccg	6360
	ctgtgcgagt	gtatgac	gggtgt	tgg	tcac	cgac	6420
	gttaggttgc	gggttact	aaacacacca	gggttgc	tgc	ccat	6480
40	ttctgggaga	gcttcttac	aggcctacc	cacat	aggccat	gtccc	6540
	aagcaggcag	gagaca	cccttac	gtac	atgg	gact	6600
	gctcaggc	cacccatc	gtgg	gatgttgc	gttcc	cgact	6660
	acgtgcac	ggcaac	ctgtgt	aggctgg	ccgtt	cgagg	6720
	accacacacc	ccat	at	ccat	gtgt	gggt	6780
	acgagcac	gggtgtt	aggcgg	ctagc	ccgc	ttgc	6840
45	acaggcagc	tgttatt	ggcagg	at	tttgc	gac	6900
	gacagggaa	tcc	ggat	ttcgat	gatgttgc	aaaa	6960
	tacatcgaa	aggaa	gtc	ggat	ggat	ccat	7020
	caaacagcc	caac	ggagg	gtct	ggat	tttgc	7080
	atcgaagct	tctgg	gaa	catat	gtgt	atatt	7140
	ggcttgc	cttgc	caac	cccg	atagc	tgtgg	7200
50	atcaccagcc	cg	cc	at	ccat	cacag	7260
	gccc	ttgt	cc	cc	ccat	ccat	7320
	gccc	gt	cc	cc	ccat	ccat	7380
	gccc	cc	cc	cc	ccat	ccat	7440
	gaggac	ttaa	cc	cc	ccat	ccat	7500

	gtgtgcgcag	cgatactgcg	tcggcacgtg	ggcccagggg	agggggctgt	gcagtggatg	7560
5	aaccggctga	tagcgttgc	ttcgcggggt	aaccacgtct	cccccacca	ctatgtgcct	7620
	gagagcgcacg	ctgcagcacg	tgtcactcg	atccctctct	gtcttaccat	cactcagctg	7680
	ctgaagaggc	ttcaccagtg	gatcaacgc	gactgttcca	cggcatgtc	cggctcggtg	7740
	ctaagagatg	tttgggatg	gatatgcacg	gtgttgactg	atttcaagac	ctggctccag	7800
	tccaaagctcc	tgccgcgatt	gccgggagtc	cccttcttct	catgtcaacg	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgco	catgtggggc	acagatcacc	7920
	ggacatgtga	aaaacggttc	catgaggatc	gtggggctta	ggacctgtag	taaacacgtgg	7980
10	catggaacat	tccccattaa	cgcgtacacc	acggggccct	gcacgcctc	cccgccgcca	8040
	aattattctt	ggggcgtgtg	gccccgtggct	gtcgaggagt	acgtggaggt	tacgcgggtg	8100
	ggggatttcc	actacgtgac	gggcgttggcc	actgcacgcg	taaagtggcc	gtgtcagggt	8160
	ccggcccccg	aatttttcac	agaagtggat	gggggtcggt	tgcacaggt	cgctccagcg	8220
	tgcaaaacccc	tcctacggga	ggaggtcaca	ttccctgggtc	ggctcaatca	atacctggtt	8280
	gggtcacage	teccatgco	gccccgaaaccg	gtgttagcag	tgcctactc	catgttcacc	8340
15	gacccttccc	acattacggc	ggagacggct	aagcgttaggc	tggccagggg	atcteetccc	8400
	cccttggcca	getcatcage	tagccagctg	tctgcgectt	ccttgaagge	aacatgcact	8460
	acccgtcatg	actccccgg	cgctgaccc	atcgaggcca	accttctgt	cgccgcaggag	8520
	atggcgccga	acatcaccc	cgtggagtca	aaaaataaagg	tagtaattt	ggacttcttc	8580
	gagccgcctc	aaggcgagga	ggatgagagg	gaatgttcc	ttccgggga	gatctgcggg	8640
	aggtccagga	aattttctcg	agcgtatccc	atatgggcac	gccccggata	caacecttca	8700
20	ctgttagagt	ccttgcggaa	cccggtactc	gttccctccag	tgttacacgg	gtgtccattt	8760
	ccgcctgcca	aggcccctcc	gataccaccc	tcacggagga	agaggacgt	tgtccctgtca	8820
	gaatctaccg	tgttttctgc	tttggggag	cttcgcacag	agaccttcgg	cagctccgaa	8880
	tcgtcgccg	tcgacagcgg	cacggcaacg	gcttctctg	accagccctc	cgacgcacggc	8940
	gacgcgggat	ccgcacgttga	gtcgtaactc	tccatgcccc	cccttgggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggtct	acgtttaagcg	aggaggctag	tgaggacgtc	9060
25	gtctgtctgt	cgatgttca	cacatggaca	ggcccccgt	tcacgcctg	cgctgcggag	9120
	gaaacaaggc	tggccatcaa	tgcactggc	aacttcttgc	tccgttacca	caacttggtc	9180
	tatgtacaa	catctcgca	cgcaagcctg	cggcagaaga	aggtcacett	tgacagactg	9240
	caggctctgg	acgaccacta	ccgggacgtg	ctcaaggaga	tgaaggcgaa	ggcgccacaca	9300
	gttaaggctt	aacttctata	cgtggaggaa	gctttaagc	tgacgcctcc	acatttgcgc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gttccggaaacc	tatccagcaa	ggccgttaac	9420
30	cacatccgct	ccgtgtggaa	ggactgtcg	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggca	aaaatgggt	tttctcgctc	caaccagaga	agggggggcc	caagccagct	9540
	cgccttatcg	tatccccaga	tttgggggtt	cgtgtgtcg	agaaaaatggc	cctttacgt	9600
	gtgtctcca	ccctccctca	ggccgtgtat	ggtctttcat	acggatttca	atactctct	9660
	ggacagcgccg	tcgagttct	ggtaatgtcc	tggaaagcg	agaaatggcc	tatgggctt	9720
	gcatatgaca	cccgctgtt	tgactcaacg	gtcactgaga	atgacatccg	tgttggaggag	9780
35	tcaatctacc	aatgttgtga	tttggccccc	gaagccagac	aggccataag	gtcgctcaca	9840
	gagcggcttt	acatcggggg	ccccctgact	aattctaaag	ggcagaactg	cggctatcgc	9900
	cgggtcccg	cgagcggtt	actgacgacc	agctcggtt	atacccttac	atgttacttg	9960
	aagcccgctg	cggccgtgt	agetcgcaag	ctccaggact	geacgatgt	cgtatgcgg	10020
	gacgaccctg	tcgttatctg	tggaaagcg	gggacccaaag	aggacgaggc	gagccctacgg	10080
	gccttcacgg	aggctatgtac	tagatactct	gccccccctg	gggacccccc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgtcc	tccaatgtgt	cagtcgcgca	cgatgcac	10200
	ggcaaaaggg	tgtactatct	cacccgtgac	ccccaccacc	cccttgcgcg	ggctgcgtgg	10260
	gagacagcta	gacacactcc	agtcaattcc	tggctaggca	acatcatcat	gtatgcgc	10320
	accttgcggg	caaggatgt	cctgtatgact	cattttcttct	ccatcccttct	agctcaggaa	10380
	caacttggaa	aagccctaga	ttgtcagatc	tcggggccct	gttacttcat	tgaccactt	10440
	gacccatcc	agatcattca	acgactccat	ggcccttagcg	cattttcat	ccatgttac	10500
45	tctccagggt	agatcaatag	ggtgggttca	tgccttcagg	aacttgggtt	accggccctt	10560
	cgagtctgga	gacatcgggc	cagaagtgtc	cgcgcgttaggc	tactgtccca	ggggggggagg	10620
	gctgccactt	gtggcaagta	ccttcttcaac	tgggcagtaa	ggaccaagct	caaactcact	10680
	ccaatcccg	ctgcgtccca	gttggattta	tccagctgtt	tcgttgcgtt	ttacagcg	10740
	ggagacatata	atcacagct	gtctcggtcc	cgaccccgct	ggttcatgt	gtgcctactc	10800
50	ctactttctg	taggggttag	catctatcta	ctccccaacc	gatgaacggg	gagctaaaca	10860
	ctccaggcc	ataggccatc	ctgttttttt	cccttttttt	tttttttttt	tttttttttt	10920
	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	10980
	tgggtccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040
	agagagtgtct	gatactggcc	tctctgcaga	tcaagt			11076

<210> 10
<211> 8001
<212> DNA
<213> Hepatitis C Virus

5 <400> 10
 gccagcccccc gattgggggc gacactccac catagatcac tccctgtga ggaactactg 60
 tttcacgca gaaagcgctc agccatggcg ttagtatag tgcgtgcag cctccaggac 120
 ccccoctccc gggagagccg tagtgctcg cggaaaccgg gatgacaccg gaattgcac 180
 gacgaccggg tccttctt gatacccg ctaatgcct ggagatttg gctgtcccc 240
 gcgagactgc tagccgagta gtgtggtc gcaaaggcc ttgtggact gcctgatagg 300
 gtgcttgcgaa tgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
 ctcaaaaaaa aaccaaaactg aacccaca ggcgcgcatt gattgaacaa gatggatgc 420
 aecagggttc tccggccgt tgggtggaga gctattcg statgactg gcaacaacaga 480
 caatcgctg ctctgatgcc gccgtgtcc gctgtcagc gcaggggcgc cgggtttcc 540
 ttgtcaagac cgacctgtcc ggtgcctga atgaaactgca ggacgaggca ggcgggtat 600
 cgtggctggc cacgacggg gttccttgcg cagctgtct gcacgttgc actgaaggcg 660
 gaaggactg gtcgtatgg ggcgaatgc cggggcagg ttcctgtca ttcacccctt 720
 ctctcgccgaa gaaatgtcc atcatggct atgcaatcg ggggtgtcat acggttgc 780
 cggctacccg cccatcgac caccaagcgaa acatcgcat cgagcgagca cgtactcg 840
 tggaaagccgg tcttgcgtat caggatgatc tggacgaa gatcagggg ctcggcccg 900
 cccaactgtt egccaggctc aaggcgcgcg tggccgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg ccgtttttctt ggattcatcg 1020
 actgtggccg gctgggtgt gccggccgt atcaggatcg acgtgttgtt acccggtata 1080
 ttgtcttgcgaa gttggccgc gaatggctg accgtttctt cgtgtttttc ggtatcgcc 1140
 ctcccgattc gcagccgcatt gccttcttgc gccttcttgc cgagggtttc tgatgttttt 1200
 cagaccacaa cggtttccctt cttagggat caattccgc cctctccctt cccccccctt 1260
 aacgttactg gccgaagccg cttggataa ggcgggtgt ctgggttgcata tatgttattt 1320
 tcaccatata tggcgctctt tggcaatgtg aggccccggaa aacctggccc tgcgtttttt 1380
 acgagccatc ctatgggtt tttcccttc gccaaaggaa tgcagggtt gttgaatgtc 1440
 gtgaaaggaaat cagtcttgc gaaaggctt tgaagaaaaaa caacgtctgt acgcggccctt 1500
 tgcaggcgcg ggaacccccc acctggcgac aggtgccttgc gggccaaaaa gccacgtgt 1560
 taagatacacat ctgcaagggc ggcacaaccc cagtgccacg ttgtgagttt gatagttgt 1620
 gaaagagtca aatggcttc ctcgaatgtt ttcacaacgg ggctgaaggaa tgcccagaag 1680
 gtaccccccatt gtatgggtt tgatctgggg cctcgggtca catgttttac atgtgtttt 1740
 tcgaggttaa aaaacgtctt gggccccccgaa accacggggaa cttgggtttt ctttggaaaa 1800
 caccataataa ccatggccgc tattacggcc tactcccaac agacgcgagg cttacttgc 1860
 tgcatacatca cttagcttcac aggccggac aggaaccagg tcgaggggga ggtccaaatgt 1920
 gtctccacccg caacacaatc tttccctggcg acctgcgtca atggcggtgt tggactgtc 1980
 tatcatgttgc cccgtctaaa gacccttgc ggcaccaatcc cccaaatgtac 2040
 accaatgtgg accaggacat cgtcgctgg caagcgcccccc cggggccgcg ttccttgaa 2100
 ccatgcaccc gccggcgtcc ggcacccattt ttgtcacgaa ggcacccggaa tgcatttcc 2160
 gtgcggccggc gggcgacag cagggggacg ctactctcc ccaggcccgt ctccacttgc 2220
 aagggtctt cggcggtcc actgtctgc ccctcgggc acgtgtggg catctttccg 2280
 gtcggcggtt gcaacccgggg ggttgcgaa ggggtggact ttgttacccgt cgaggtctatg 2340
 ggaacacta tgcgggtccc ggtcttcacg gacaactctgt cccctccggc cgtaccggc 2400
 acatcccgagg gggccatcat acacggccctt actggtagcg gcaagacgcac taagggtccg 2460
 gtcgcgtatg caggccaaagg gtataagggtt ctgttcttgc acccggtccgt cgccgcacc 2520
 cttagtttcg gggcgatata gtcataaggca catggatcg accctaaatcat cagaatcg 2580
 gtaaggacca tcacccacggg tgccccatc acgtactcca cctatggca gtttcttgc 2640
 gacgggtgtt gctctggggg cgcctatgac atcataatat gtgtatgggtt ccactcaact 2700
 gactcgacca tctatctggg catcgacca gtcctggacc aacggcgac ggtcgagcg 2760
 cgactcgctg tgctcgccac cgtaegegtt cgggatcg tccacgtgc acatccaaac 2820
 atcgaggagg tggctctgtc cagcactggaa gaaatccccctt tttatggca gacccatcccc 2880
 atcgagacca tcaagggggg gaggcacctc attttctgccc attccaagaa gaaatgtgt 2940
 gagctcgcccg cgaactgttc cggccatcg ctaatgtcg tagcatatta cggggccctt 3000
 gatgtatcccg tcataccatc ac tagcgagac gtcattgtcg tagcaacggaa cgtctataatg 3060
 acggggcttta cccggacccgtt cggactcgatcgtt atcactgtcg atacatgtt caccacagaa 3120
 gtcgacttca gcttgcggcc gacccatcacc attgagacga cggccgtccg acaagacgc 3180
 gtgtcacgtt cgcaggccggc aggccaggact ggttagggca ggtatgggcattt acagggtttt 3240
 gtgactccag gagaacggcc ctcggcatg ttgattctt cgggttctgt cgagtgtctat 3300
 gacgcggggcgtt gtcgttggta cggactcgatc cccggccgaga cctcgttagt gttggccggct 3360
 tacctaaaca caccagggtt gcccgtctgc caggaccatc tggatgttctgtt gggagacgc 3420
 ttacccatcatc acgacccat ttttgcgtcc agactaaagca ggcaggagac 3480
 aacttccctt acctggtagc ataccaggct acgggtgtcg ccagggtctca ggctccaccc 3540

ccatcgtgg accaaatgtg gaagtgttcc atacggctaa agcctacgt gcacgggcca 3600
 acgccectgc tgtataggct gggagccgtt caaaacgagg ttactaccac acacccata 3660
 accaaataca tcatggcatg catgtcggt gacctgggg tcgtcacag caccctgggt 3720
 ctggtaggcg gagtccttagc agctctggcc gcgtattgcc tgacaacagg cagcgtggc 3780
 attgtggca ggatcatett gtccggaaag cggccatca tteccgacag ggaagtctt 3840
 taccggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
 atgcagctcg ccgaacaatt caaacagaag gcaatccggg tgctgcaac agccaccaag 3960
 caagcggagg ctgctgtcc ctgggtggaa tccaatggc ggaccatcg agccttctgg 4020
 gcgaacata tggtaattt catcagcggg atacaatatt tagcaggct gtccactctg 4080
 cctggcaacc cggcgatagc atcaactgtg gcattcacag cctcttatcac cagcccgctc 4140
 accacccaaac ataccctctt gtttaacatc ctggggggat gggtgccgc ccaacttgc 4200
 cctcccagcg ctgccttcgc ttctctggc gcccgcattc ctggagccgc tggtgcagc 4260
 ataggcctcg ggaagggtct tggtgatatt tggtcggtt atggagcagg ggtggcaggc 4320
 gcgcctgtgg cctttaaggt catagcggc gagttgcctt ccaccgagga cctggtaac 4380
 ctactccctg ctatccctc ccctggcgcctt cttagtcgtc gggtcgtgt cgacgcgata 4440
 ctgcgtcggc acgtggggcc aggggaggg gctgtcgagt ggatgaaccc gctgatagcg 4500
 ttgcgttcgc gggtaacca cgtctcccc accgcactat tgctcgagag cgacgtcgca 4560
 gcacgatca ctcagatct ctctgttcc accatactc agctgtcaa gaggcttcac 4620
 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680
 gattggatat gcacgggtt gactgattt aagacctggc tccagtccaa gctccctgccc 4740
 cgatggcgg gagtccctt ctctctcatgt caacgtggg acaaggaggat ctggggggc 4800
 gacggcatca tggaaaccac ctggccatgtt gggacacaga taccggaca tggaaaaaac 4860
 gttccatga ggatcggtt ggcttaggacc tgtagtaaca ctgtggatgg aacatcccc 4920
 20 attaacgcgt acaccacggg ccctgcacg ccctccccgg cgccaaatata ttctagggcg 4980
 ctgtggcggt taggtgtcga ggagtacgtg gagggtacgc gggtggggaa ttccactac 5040
 gtgacggcga tgaccactga caacgtaaag tgccccgtc aggttccggc ccccgaaattc 5100
 ttcacagaaag ttgatgggggt gcggttgc acgtacgtc cagcgtgcaaa accccctccct 5160
 cgggaggagg tcacattctt ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
 tgcgagccctg aaccggatgt agecaactgtc acttccatgc tcaccgaccc ctcccacatt 5280
 acggcggaga cgctaaagcg taggtgtggc aggggatctc ccccccattt ggcctactca 5340
 tcagcttagcc agctgtcgcc gccttccttgg aggcaacatc gcaactaccc tcatgactcc 5400
 ccggacgcgtc acctcatcgaa ggcacccctc ctgtggcgcc aggagatggc cgggaacatc 5460
 accccgcgtgg agtcagaaaa taaggtagt attttggact ctttcgagcc gctccaaagcg 5520
 gaggaggatg agaggggagt atccgttccg gggagatcc tgccggaggtc cagggaaattc 5580
 cctcgagcga tgcccatatggc acgtggcccg gattcaacc ctccactgtt agagtccctgg 5640
 aaggacccgg actaactgtcc tccatgttgc acgggggttgc cattggccgc tgccaaaggcc 5700
 30 cctccgatcatc caccttcacg gaggaaagggg acgggtgtcc tgcacaaatc taccgtgtct 5760
 tctgccttgg cggagctcgcc cacagagacc ttccggcagct ccgaatcgcc gggcgtcgac 5820
 agccggcacgg caacggccctc tcctgaccag ccctccgcacg acggcgcacgc gggatccgac 5880
 gttgagtcgt actccctcatatggcccccctt gggggggatcc cggggatcc cgatctcagc 5940
 gacgggttccg ggtctacgtt aaggaggagg gctagtgggg acgtcgctt ctgcgtcgatg 6000
 tcctacacat ggacaggccg cctgtatcgcg ccatgtcgccg cggagggaaac caagctgccc 6060
 atcaatgcac tgagcaactc tttgtccgtt caccacaact tggtgtatgc tacaacatct 6120
 cgcagcgcacaa acctgcggca gaagaagggtc acctttgaca gactgcagggt cctggacgc 6180
 cactcgccgg acgtgtccaa ggagatggg gccaaggccgtt ccacagttaa ggctaaactt 6240
 ctatccgtgg aaggacgttgc taatgttgc ccccccattt cggccatgc taaatttggc 6300
 tatggggcaaa aggacgtccg gaaacctatcc agcaaggccg ttaaccacat cgcctccgt 6360
 tggaaaggact tgctggaaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
 gagggtttctt gcgtccaaacc agagaagggg gggccgcacgc cagctcgccct tatcgatttc 6480
 40 ccagatttgg ggggtcggtt gtgcggaaaaa atggccctttt acgtatgttgc ctccaccctc 6540
 cctcaggcccg tgatgggttgc ttccatagggatc ttccaaatactt ctctggaca ggggtcgag 6600
 ttccctggta atgcctggaa agcgaagaaaa tggccctatgg gtttcgcata tgacacccgc 6660
 tgttttgact caacgttcac tgagaatgtac atccgtgttgg aggagtcaat ctaccaatgt 6720
 tgcgtactgg ccccccggc acgtggccatc ataagggtcgc tcacagagcg gctttacate 6780
 gggggcccccc tgcataatcc taaaggggcag aactggccgtt atccggccgtg cccgcgcgagc 6840
 ggtgtactgg cgaccatgtt cggatataactt ctcacatgtt acttggggcccg cgctgcggcc 6900
 tgcgtactggt cgaagctcaaa ggactgcacg atgtctcgat gcccggacgc ccttgcgtt 6960
 atctgtggaaa ggcgggggac ccaagaggac gaggccggcc tacgggcctt cacggaggct 7020
 atgacttagt actctgccttcc cccctggggac ccccccacaaac cagaatacga cttggagttg 7080
 ataacatcat gtcctccaaat tgcgtactggt cgcgcacgtatc catctggcaaa aagggtgtac 7140
 tatctccatcc gtcggccatcc ccccccattt gggccggccgtt cttggggagac agctagacac 7200
 actccagtcatc attcctggatc agggaaacatc atcatgtatc cggccacccctt gtggcaagg 7260
 atgatctgttgc tgactcattt ctcttcatc ttctctgttgc tccatgttgc cacttgaccc acctcagatc 7320
 ctagattgttgc agatctacgg ggcctgttac tccatgttgc cacttgaccc acctcagatc 7380
 attcaacgcac tccatggccct tagcgttgcattt tcaactccatca gttactctcc aggtgagatc 7440
 aatagggtgg ctcatgttgc caggaaactt ggggttacccgc ctttgcgtt cttggagacat 7500

EP 1 043 399 A2

cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
aagtacctt tcaactggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
5 tcccagttgg atttatccag ctggttcggt gctggttaca gcgggggaga catatatcac 7680
agcctgtctc gtgcccgacc ccgctggtc atgtggtgcc tactcctact ttctgttaggg 7740
gttaggcattt atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800
ccatcctgtt ttttccctt tttttttttt tttttttttt tttttttttt tttttttttt 7860
ttttctcctt tttttttctt ctttttttcc ttttcttcc tttggtggtt ccatcttagc 7920
cctagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980
10 tggcctctct gcagatcaag t 8001

15

20

25

30

35

40

45

50

55

<210> 11
 <211> 11076
 <212> DNA
 <213> Hepatitis C Virus

<400> 11
 gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tcttcacgca gaaagcgct agccatggcg ttatgttag tgctgtcgac cctccaggac 120
 cccccctccc gggagagcca tagtggctcg cgaaaccggg gagaatccacgg gaattgcacg 180
 10 gacgaccggg tccttcttgc gatcaaccgg ctcatacgct ggagatttgg gcgtgcggcc 240
 gcgagactgc tagccgagta gtgtgggtc gcgaaaaggcc ttgtgttact gcctgtatagg 300
 gtgcttgca gtcgccccggg aggtctcgta gaccgtgcac catgagacg aatcctaacc 360
 ctcaaagaaa aaccaaacgt aacccaacg ggcgcgccat gattgaacaa gatggattgc 420
 15 acgcagggtc tccggccgct tgggggaga ggttatccgg ctatgactgg gcacaacaga 480
 caatcggtcg ctctgtatggc ggcgtgtcc ggctgtcagc gcaaggggcgc cccgttctt 540
 ttgtcaagac cgaccgttgc ggtccctgta atgaaactgca ggacgaggca ggcgggttat 600
 cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaaggcg 660
 gaagggactg gctgtatgg ggcgaagtgc cggggcagga ttcctgtca ttcacettg 720
 ctctgcgcg aaaaatgtcc atcatggct atgaaactgca ggccgggtcat acgctgtatc 780
 20 cggttacatcg cccatcgac caccacgca aacatcgca cgagcgagca cgtactcgga 840
 tggaaaggccg tcttgcgtat caggatgtc tggcgaaga gatcagggg ctcgcgcggc 900
 cggactgtt cgccaggctc aaggccgcga tgcccgacgg cgaggatctc gtcgtgaccc 960
 atggcgatgc ctgttgcg aatatcatgg tggaaaatgg cgcgtttctt ggattcatcg 1020
 actgtggccg gctgggtgtc gccggccgtc atcaggacat agcgttggtc acccgtgata 1080
 ttgtcttgcgaa gttggccgc gatggctgtc acgctttctt cgtgtttatgg ggtatccggc 1140
 ctcccgatttgc cagccgtatc gccttctatc gccttcttgc cagtttttc tgatgtttaaa 1200
 25 cagaccacaa cgggttccctt ctggggat ctaatccggcc cctctccctt cccccccctt 1260
 aacgttaactg gccgaagccg ctttggataa ggccgggtgtc ctgggttgc tatgttattt 1320
 tccaccatat tggcttgcgtt tggcaatgtg agggccggaa aacctggccc tggcttctt 1380
 acgagcattc ctagggggtc ttcccttcgca gcaaaaggaa tgcacgggtt gttgaatgtc 1440
 30 gtgaaggaaag ctttgcgttgc ggaaggttt tgaagacatc aacatcgatc agcggaccct 1500
 tgcaggcgc ggaacccccc acctggcgac aggtgccttc gggccaaaa gccacgtgt 1560
 taagatacac ctgcaaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagtgt 1620
 gaaaagagtca aatggcttc ctcaacgca ttcacacaagg ggctgaagga tgcccagaag 1680
 gtaccccaattt gtatgggtatc tgatgggggg cctcggtgca catgttttac atgtgtttag 1740
 tcgaggatcaa aaaacgttca gggcccccggc accacggggg cttgggtttc ctttggaaaaa 1800
 35 cagataataa ccatggggac gaaatctaaa cctcaaaaggaa aaaccaaaacg taacaccaac 1860
 cggcccccac aggacgtcaa gttccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
 ttggccgcga gggggcccaag gttgggtgtc cgcgcgacta ggaagacttc cgagcggtcg 1980
 caacctcggtg gaaaggcaca acctatcccc aaggctcgcc agcccgaggg tagggccctgg 2040
 gtcagccgc ggttacccctg gccccttatc ggcacatggg gcttgggtgg ggcaggatgg 2100
 ctctgttgcac ccgggttgcgc tcggccatgt tggggcccaacggccccc gctgttaggtcg 2160
 40 cgcacatttgg gtaaggatcat cgataccctc acgtggggct tccgcgtatc catgggttac 2220
 attcgcgtcg tcggcgcccc ccttagggggc gtcgcgggg ccctggcgca tggcgccgg 2280
 gttctggagg acggcgatc ctatgcaaca gggaaatctgc cccgttgc cttttctatc 2340
 ttccctttgg ctgtgtgtc ctgtttgtacc atcccgatc cccgttcatatc agtgcgcac 2400
 gtatccggag ttttgcgtatc cacaacgac tgcgttcaacg caagcatgtt gatgtgaggca 2460
 gcgacatcatc ttcgttgcac cccgggtgc tgccctgtc ttccggggaaa caactccctt 2520
 cgcgtgtggg tagcgctcac tccacacgtc gggccaggaa acgctagcg ccccaactacg 2580
 acgataacgac gccatgtcgaa ttgtcggtt gggccggcgtc ctctctgtc cgcgtatgtac 2640
 45 gtggggatc tctcggtatc tggtttccctc gtcgcggccatc tggttgcaccc ttcgcgttgc 2700
 cggcgcgaga cagtagacca ctgcgtatc tcaatataatc cccggccacgt gacagggtcac 2760
 cgtatggctt gggatgtatc gatgtactgg tcacccatcg cagccctatgt ggtatcgcc 2820
 ttactccggatcc cccacaaatc tgctgtgtatc atgttgggggg gggcccatgg gggaggctca 2880
 gcccccttgc cctactattc catgggggg aactgggtca aggttctgtat tggatgtatc 2940
 ctctttcccg gcggttgcgg gggaaacctat gtgcacagggg ggacgtatggc caaaaacacc 3000
 ctccggatata cgttccctttt ttcacccggg tcaatccaga aaatcccgatc tgtaaaacacc 3060
 aacggcgatc ggcacatcaa caggactggc ctgaaactgca atgactccct caacactggg 3120
 ttccctgtcg cgtgttgc tgcgtacaaatc ttcacccatcg ctggatgtccc agagcgatcg 3180
 gccagctgca gccccatcgca cgcgttgcgtc caggggtggg ggcccatcac ttacaatgtag 3240
 tcacacagct cggaccaggag gccttattgt tggcactacg caccggccggc gtcgtgtatc 3300
 gtacccgggg cgcgtgtcg tggccatgtc tactgttca ccccaagccc tggatgtatc 3360
 50 gggacgaccg acgggttccg cgttccatcg tacatgggggg gggagaatga gacggacgtg 3420
 ctgttccatcg tcaacacgca ggcgcggccaa ggcacatgttgc tggctgtatc agatgtatc 3480
 agcaactgggt tcaccaaaatc gtcgcggggc ccccccgttgc acatcggggg gatcgccaat 3540

	aaaacccctga	cctgccccac	ggactgcttc	cggaagcacc	ccgaggccac	ttacaccaag	3600
5	tgtggccgg	ggccttgggt	gacaccaga	tgcttggcc	actacccata	caggctttgg	3660
	cactaccct	gcactgtcaa	cttaccatc	ttcaagggtta	ggatgtacgt	ggggggagtg	3720
	gagcacagggc	togaagccqc	atgcaattgg	actcgaggag	agcggtgtaa	cctggaggac	3780
	agggacagat	cagagcttag	cccgtctgt	ctgtctacaa	cgagtgccgca	ggtattgccc	3840
	tgttccctca	ccaccctacc	ggctctgtc	actggttga	tccatctcca	tcagaacgtc	3900
	gtggacgtac	aataccctgt	cggtataggg	tcggcggttgc	tctctttgc	aatcaaatgg	3960
	gagtagtgtcc	tgttgctctt	ccttcttc	ggggacgcgc	gggtctgtgc	ctgtttgtgg	4020
10	atgatgtgc	tgatagctca	agctgaggcc	gccttagaga	acctgtgtgt	cctcaacgcg	4080
	gcateccgtgg	ccggggcgca	tggcattctc	tccttccctcg	tgttcttctg	tgctgcctgg	4140
	tacatcaagg	gcaggcgtgt	ccctggggcg	gcataatgcc	tcacggcggt	atggccgccta	4200
	ctcctgtcc	tgtggcttc	accaccacga	gcatacgca	tgaccgggaa	gatggcagca	4260
	tctgtccggag	ggcggttgc	cgttaggtctg	atacttgc	ccttgcacc	gcataataag	4320
	ctgttccctcg	ctaggctcat	atggtgggtt	caatattttt	tcaccaggggc	cgaggcacac	4380
15	ttgcaagtgt	ggatcccccc	cctcaacgtt	cgggggggcc	gcgtatggcg	catccctctc	4440
	acgtgcgcga	tccacccaga	gtaatctttt	accatcacca	aaatcttgc	cgcctatactc	4500
	ggtccactca	tggtgcctc	ggctgggtata	accaaagtgc	cgtactctgt	ggcgccacac	4560
	gggcttcatc	gtgcgtcat	gtgggtcg	aaagggtctg	gggggtcatta	tgtccaaatg	4620
	gctctcatga	agttggccgc	actgacaggt	acgtacgtt	atgaccatct	caccccaactg	4680
	cgggacttggg	cccacgcggg	cctacgagac	cttgcgggtgg	cagttgagcc	cgtcgcttc	4740
	tctgtatgg	agaccaagg	tatcacttgc	ggggcagaca	ccggccgtgt	tggggacatc	4800
20	atcttggcc	tgcctgtc	ccggccgg	ggggggggaa	tatcat	accggcagac	4860
	agccttgaag	ggcagggtgt	gggggggg	gggggggg	gggggggg	ccaaacagacg	4920
	cgaggcctac	ttggctgtc	catcacttgc	ctcacaggcc	ggggcaggaa	ccaggtcgag	4980
	ggggagggtcc	aagtggctc	cacccgaaca	caatcttcc	tggcaccatc	cgtcaatggc	5040
	gtgtgttgg	ctgttctatc	tggtgcggc	tcaaaagacc	ttggccggccc	aaaggggccca	5100
	atcaccctaa	tgtacaccaa	tgtggaccag	gacccgtcg	gtggccaa	ggcccccggg	5160
	gcccgttctt	tgacaccatc	cacccgtggc	agctggacc	tttacttggt	cacgaggcat	5220
25	gccgatgtca	ttccgggtcg	ccgggggggc	gacaggcagg	ggggcctact	ctcccccaagg	5280
	cccgctctct	acttgaagg	cttccggggc	gttccactgc	tctggccctc	ggggcacgct	5340
	gtgggcattc	ttccgggtcg	ctgtgtcacc	cgagggttgc	cgagggttgc	ggactttgt	5400
	cccgctgagt	ctatggaa	cactatcg	tcccccgtct	tcacggacaa	ctcgccccct	5460
	ccggccgtac	cgcagacatt	ccagggtggc	catctacacg	ccctctactgg	tagcggcaag	5520
	agcaactaagg	tgccggctgc	gtatgcaggc	caagggtata	aggtgtttgt	cctgaacccg	5580
	tccgtcgccg	ccacccctagg	tttccggggc	tataatgtcta	aggccacatgg	tatcgaccct	5640
30	aacatcagaa	tccgggttga	gaccatcacc	acgggtggcc	ccatcactg	ctccacccat	5700
	ggcaagttt	ttggccgacgg	tgggtgtct	ggggggccct	atgacatcat	aatatgtat	5760
	gagtgcact	caactgactc	gaccacttac	ctgggcatcg	gcacagtctt	ggaccaagcg	5820
	gagacggctg	gagcgcgact	cgtcggtctc	gcccggctt	cgccctccggg	atcggtcacc	5880
	gtgccacatc	caaacatcg	ggagggtgg	ctgtccacga	ctggagaaat	cccttttat	5940
	ggcaaaggca	tcccccatacg	gaccatcacc	gggggggggg	acccatctt	ctggcattcc	6000
35	aagaagaaat	gtgtatgg	cgccgcgaag	cttccggggc	tgggactcaa	tgctgttagca	6060
	tattaccggg	gccttggatgt	atccgtcata	ccaaactagcg	gagacgtcat	tgctgttagca	6120
	acggacgctc	taatgacggg	tttacccggc	gacttcgact	cagtgtatcg	ctgaataaca	6180
	tgtgtcaccc	agacagtctg	tttcggcttc	gaccggac	taccatcg	gacgacgacc	6240
	gtgccacaa	acggcggtgc	acgtcgcc	cgggaggagg	ggacttggtag	ggcaggatg	6300
	ggcatttaca	ggtttggatc	tccaggagaa	cgccctcg	gcatgttgc	ttccctgggt	6360
	ctgtgcgagt	gttatgtac	ggggctgtgt	tggta	ccacggc	cgagacctca	6420
	gttaggttgc	gggtttactt	aaacacacca	gggttggcc	tctggccagg	ccatctggag	6480
40	ttctggggaa	gcgttccatc	agggcttacc	cacatagacg	ccatcttctt	gtcccagact	6540
	aaggcaggcg	gagacaaatc	cccccatacg	gtacatcacc	aggctacgtt	gtgcggcagg	6600
	gctcaggcgc	cacccatccatc	gtgggaccaa	atgttggaa	gttcatacg	gttccatc	6660
	acgctgcacg	ggccaaacgc	cctgtgtat	aggctgggg	ccgttcaaaa	cgaggttact	6720
	accacacaccc	ccataacccaa	atacatcatg	gcatgtatgt	ccggctgac	ggagggtcg	6780
	acgagcacct	gggtgtgtgt	aggcggaggc	ctagcagtc	tggccggcgta	ttgcctgaca	6840
45	acaggcaggcg	tggtcatttg	ggggcggatc	atcttgc	ggaaaggccgc	catattccc	6900
	gacagggaag	tcccttaccg	ggagttcgat	gagatggaa	agtgccctc	acacccctt	6960
	tacatcgaa	aggaaatgt	gtcgccgaa	caattcaaa	agaaggcaat	cggtttgt	7020
	caaacagcca	ccaaacaa	ggaggctgt	gtccctgtgg	tggaaatccaa	gtggccgacc	7080
	atcgaaggct	tctgggggaa	gcatatgtgg	aatttcatca	gggggatata	atatttagca	7140
	ggcttgc	ctctgcctgg	caaccccgcc	atagcatc	tatggcatt	cacagccct	7200
50	atcaccagcc	cgctcaccac	ccaaatacc	cttcgttta	acatcttgg	gggtatgggt	7260
	gccgccccaa	ttgttccctcc	cagcgtgt	tctgtttcg	taggcggcc	catcgctg	7320
	gcggctgttg	gcacgtatgg	ccttggaaag	gtgttgg	atattttggc	aggttatg	7380
	gcaggggtgg	caggcgcgt	cgtggccctt	aaggctatg	gcccggcgat	gcccctccacc	7440
	gaggacacttgg	ttaacact	ccctgtatc	cttccctctg	gcccctatgt	cgtcggggtc	7500

	gtgtgcgcag	cgatactgcg	tcggcacgtg	ggccccagggg	agggggctgt	gcagtggatg	7560
5	aaccggctga	tagcgttgc	ttcgcgggggt	aaccacgtct	cccccaecga	ctatgtgcct	7620
	gagagcgcacg	ctgcagcaca	agtcaactcg	actctctcta	gtcttacca	cactcagctg	7680
	ctgaagaggc	ttcaccagt	gatcaacgag	gactgtccca	cgccatgtc	cggtctgtgg	7740
	ctaagagatg	tttgggattt	gatatgcacg	gtgttgactg	atttcaaga	ctggctccag	7800
10	tccaaagctcc	tgccgcgatt	gccggggatc	cccttcttct	catgtcaac	tgggtacaag	7860
	ggagtctggc	ggggcgacgg	catcatgcaa	accacctgcc	catgtggagc	acagatcacc	7920
	ggacatgtga	aaaaccggtt	catgaggata	gtggggcta	ggacactgtag	taaacacgtgg	7980
	catggaacat	tccccattaa	cgcgtacacc	acggggcccct	gcacgcctc	cccgccgcca	8040
	aattatttcta	ggggcgtgt	gccccgttaggt	gtcgaggagt	acgtggaggt	tacgcgggtg	8100
	ggggattttc	actacgtgac	gggcgtatgacc	actgacaacg	taaaagtggcc	gtgtcagggt	8160
	ccggcccccg	aatttttcac	agaagtggat	gggggtcggt	tgcacaggt	cgctccagcg	8220
15	tgc当地	tcctacgggg	ggagggtaca	ttctctggtcg	ggctcaatca	atactgggtt	8280
	gggtcacagc	tcccatgcga	gcctgaaccg	gatgttagcag	tgctcaactc	catgtcacc	8340
	gacccttccc	acattacggc	ggagacggct	aagcgttaggc	tggccagggg	atctccccc	8400
	cccttggcca	gctcatcagc	tagccagctg	tctgcgcctt	ccttgaaggc	aacatgcact	8460
	accggctcatg	actccccgg	cgctgaccc	atcgaggcc	acctcctgt	gcggcaggag	8520
	atgggcggga	acatcaccc	cgtggagtca	aaaaataagg	tagtaatttt	ggactctttc	8580
	gaggcgtcc	aaggcgagg	ggatgagagg	ggagtatccg	ttccgggg	gatccctggg	8640
	aggccaggaa	aattttctcg	acgcgtggcc	atatgggcac	gccccgatta	caaccctcca	8700
20	ctgttagagt	ccttggaaagg	cccgactac	gttccctccag	tggtaacacgg	gtgtccattt	8760
	ccgcctgcca	aggccccctt	gataccac	tcacggagga	agaggacgg	tgtcctgtca	8820
	gaatctaccg	tgttttctgc	tttggggag	ctggccacag	agaccttcgg	cagetccgaa	8880
	tcgtcgccg	tcgacagcgg	cacggcaac	gcccctctcg	accagecc	cgacgcggc	8940
	gacgcgggat	ccgacgttga	gtcgtaactc	tccatgc	cccttgggg	ggagccgggg	9000
	gatcccgatc	tcagcgacgg	gtcttggt	acgtaaagcg	aggaggctag	tgaggacgtc	9060
25	gtctgttgt	cgatgtcct	cacatggaca	ggccccctga	teaegecat	cgctcgagg	9120
	gaaatccaagc	tgcccatcaa	tgactgtgc	aacttttgc	tccgtccacca	caacttggc	9180
	tatgtctacaa	catctcgac	cgcaaaacct	cggtcagaaga	aggtcacett	tgacagactg	9240
	caggtcttgg	acgaccacta	ccgggacgt	ctcaaggaga	tgaaggcgaa	ggcgccaca	9300
	gttaaggcta	aacttctatc	cgtggaggaa	gcctgtaa	tgacgcccc	acattcggcc	9360
	agatctaaat	ttggctatgg	ggcaaaaggac	gttccggaa	tatccagaa	ggccgttaac	9420
30	cacatccgt	ccgtgtggaa	ggactgtgc	gaagacactg	agacaccaat	tgacaccacc	9480
	atcatggcaa	aaaatgggt	tttctcgct	caaccagaga	agggggccg	caagccaget	9540
	cgccttatcg	tattcccca	tttgggggtt	cgtgtgtcg	agaaaaatggc	cctttacgat	9600
	gtgtctcca	ccctccctca	ggccgtgat	ggcttctcat	acggatttca	atactctct	9660
	ggacagccggg	tcgagttcct	ggtgaatgcc	tggaaagcga	agaaatgccc	tatgggctt	9720
35	gcatatgaca	cccgctgttt	tgactcaac	gtcaactgaga	atgacatccg	tgttgaggag	9780
	tcaatctacc	aatgttgt	tttggccccc	gaagccagac	aggccataa	gtcgctcaca	9840
	gagcggtttt	acatcggggg	ccccctgtact	aattctaaag	ggcagaact	cggctatcgc	9900
	cggccggcgc	cgagcggt	actgacgacc	actgcggta	ataccctc	atgttactt	9960
	aaggccgtg	cggccgtgt	agtcgcaag	cttcaggact	gcacgtat	cgatgcgg	10020
	gacgacctt	tcgttatct	tgaaagcgc	gggacccaag	aggacgaggc	gagccctacgg	10080
	gccttcacgg	aggctatgac	tagatactc	gccccccctg	gggacccccc	caaaccagaa	10140
40	tacgacttgg	agttgataac	atcatgtcc	tccaatgtgt	cagtcgc	cgatgcac	10200
	ggccaaaagg	tgtactatct	caccgtgac	cccaccaccc	cccttgcgc	ggctgcgtgg	10260
	gagacagcta	gacacactc	agtcaattcc	tggcttaggc	acatcat	gtatgc	10320
	accttgtgg	caaggatgt	cctgtatgact	catttcttct	ccatcctt	agctcaggaa	10380
	caacttggaa	aaggccctag	ttgtcagatc	tcaggggct	gttacttcat	tgagccactt	10440
	gaccctaccc	agatcattca	acgactccat	ggccatgtcg	cattttcact	ccatagttac	10500
45	tctccagggt	agatcaatag	ggtggctca	tgcctcagga	aacttgggtt	accgcctt	10560
	cgagtcttgc	gacatcgggc	cagaagtgtc	cgcgcttagc	tactgtcc	gggggggg	10620
	ggtccactt	gtggcaagta	ccttctcaac	tggcagtaa	ggaccaag	caaactcact	10680
	ccaatcccg	ctgcgtccca	gttggattt	tccagctgt	tcgttgc	ttacagcggg	10740
	ggagacat	atcacagcc	gtctcggtc	cgaccccgct	ggttcatgt	gtgcctactc	10800
50	ctactttctg	tagggtagg	catctatcta	ctcccaacc	gatgaacgg	gagctaaaca	10860
	ctccaggc	ataggccat	ctgtttttt	ccctttttt	ttttttttt	ttttttttt	10920
	ttttttttt	ttttttttt	ttttttttt	ttttttttt	ttttttttt	ttttttttt	10980
	tggctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactc	11040
	agagagtgt	gatactggcc	tctctgcaga	tcaagt			11076

Patentansprüche

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, dadurch gekennzeichnet,

5 daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

10 2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen von einer handelsüblichen Hepatomazelllinie abstammen .

15 3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

daß die Hepatomazellen aus einer Hepatoma primärzellkultur gewonnen sind.

20 4. HCV-RNA-Konstrukt, dadurch gekennzeichnet,

daß es die HCV-spezifischen RNA-Abschnitte 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3' NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.

25 5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß es eine Nukleotidsequenz gemäß einem der Sequenzprotokolle SEQ ID NO: 1 bis SEQ ID NO: 11 umfaßt.

30 6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet,

daß die 3' NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

35 (a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT
TTAGCTTTTTTTTTCTTTTTTGAGAGAGAGTCTCACTCTG
TTGCCAGACTGGAGT

40

45

50

55

- (b) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTC TTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCCAGACT GGAGC

(c) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTAATCTT TTTTTTTCT TTTTTTTGA GAGAGAGAGT
CTCACTCTGT TGCCCAGACT GCAGC

(d) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTAGTC TTTTTTTTT TCTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT

(e) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT TCTTTTTTT TGAGAGAGAG
AGTCTCACTC TGTTGCCAG ACTGGAGT

(f) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT TCTTTTTTT TTGAGAGAGA
GAGTCTCACT CTGTTGCCCA GACTGGAGT

(g) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTAGTCT TTTTTTTTT CTTTTTTTT GAGAGAGAGA
GTCTCACTCT GTTGCCCAGA CTGGAGT

(h) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAAAT CTTTTTTTT TTTTCCCTT TTTTGAGAGA
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T

(i) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT
TTTTTTAAC TTTTTTTTT TTTTCTTTT TTTTGAGAG
AGAGAGTCTC ACTCTGTTGC CCAGACTGGAG GT

7. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,

daß das selektierbare Markerogen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.

8. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 6, dadurch gekennzeichnet,

daß das selektierbare Markerogen ein Neomycinphosphotransferasegen ist.

9. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 8, dadurch gekennzeichnet,

daß das selektierbare Markergen strangabwärts der 5' NTB in die HCV-BNA integriert ist

10. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 9, dadurch gekennzeichnet

daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.

11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet,

5 daß es ein integriertes Reportergen aufweist.

12. HCV-RNA-Konstrukt nach Anspruch 11, dadurch gekennzeichnet,

10 daß das Reportergen ein Gen aus der Gruppe der Luziferasen, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), der GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen) ist.

13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 11, dadurch gekennzeichnet,

15 daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.

14. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 13, dadurch gekennzeichnet,

20 daß das Resistenzgen derart in das offene Leseraster der HCV-RNA einkloniert ist, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführbar ist.

15. HCV-RNA-Konstrukt nach einem der Ansprüche 11 bis 14, dadurch gekennzeichnet,

25 daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionsprotein exprimieren.

16. Zellkultursystem nach einem der Ansprüche 1 bis 3 dadurch gekennzeichnet,

30 daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 15 ist.

17. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet,

35 daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBl 9-13) hinterlegt sind.

18. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.

40 **19.** Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 16 bis 17 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

20. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 15 zur Herstellung einer leberzellspezifischen Genfahre für die Gentherapie .

21. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet,

45 50 daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

22. Verfahren zur Gewinnung von zellkultur-adaptierten Mutanten eines HCV-RNA-Konstrukts gemäß einem der Ansprüche 4 bis 15, wobei die Mutanten gegenüber dem HCV-RNA-Konstrukt eine erhöhte Replikationseffizienz aufweisen, dadurch gekennzeichnet,

55 daß man ein Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt mit Selektionsgen nach einem der Ansprüche 4 bis 15 ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, daß man die gewachsenen Zellklone erntet, und daß man

aus diesen Zellklonen die HCV-RNA-Konstrukte oder Teile davon isoliert.

23. Verfahren nach Anspruch 22, dadurch gekennzeichnet,

5 daß man die isolierten HCV-RNA-Konstrukte wenigstens einmal erneut passagiert, nämlich in Zellen eines Zellkultursystems nach Anspruch 1 einschleust, das dabei erhaltenes Zellkultursystem gemäß Anspruch 1, bei dem das eingeschleuste HCV-spezifische Genmaterial das isolierte HCV-RNA-Konstrukt mit Selektionsgen ist, auf/in dem dem Selektionsgen entsprechenden Selektionsmedium kultiviert, die gewachsenen Zellklone erntet und aus diesen Zellklonen die HCV-RNA-Konstrukte isoliert.

10 24. Verfahren zur Herstellung von Mutanten eines HCV-Vollängengenoms oder eines HCV-Teilgenoms oder eines beliebigen HCV-Konstrukts mit im Vergleich zu dem ursprünglichen HCV-Vollängengenom oder -Teilgenom oder HCV-RNA-Konstrukt erhöhter Replikationseffizienz, dadurch gekennzeichnet,

15 daß man mit einem Verfahren nach Anspruch 22 oder 23 eine zellkultur-adaptierte Mutante eines HCV-RNA-Konstrukts herstellt und isoliert,
daß man die Nukleotid- und Aminosäuresequenz dieser Mutante bestimmt und durch Vergleich mit der Nukleotid- und Aminosäuresequenz des ursprünglichen HCV-RNA-Konstrukts die Art, Anzahl und Positionen der Nukleotid- und Aminosäuremutationen bestimmt,
20 und daß man diese Mutationen entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten, in ein (isoliertes) HCV-Vollängengenom oder ein HCV-Teilgenom oder ein beliebiges HCV-RNA-Konstrukt einführt.

25. Zellkultur-adaptiertes HCV-RNA-Konstrukt mit hoher Replikationseffizienz, dadurch gekennzeichnet,

25 daß es durch Nukleotid- und/oder Aminosäure-Mutationen von einem HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15 ableitbar ist und daß es mit einem Verfahren nach einem der Ansprüche 22 bis 24 erhältlich ist.

30 26. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25, dadurch gekennzeichnet,

daß es einen oder mehrere der nachfolgend aufgeführten Aminosäureaustausche aufweist, nämlich 1283 arg -> gly und/oder 1383 glu -> ala und/oder 1577 lys -> arg und/oder 1609 lys -> glu und/oder 1936 pro -> ser und/oder 2163 glu -> gly und/oder 2330 lys -> glu und/oder 2442 ile -> val aufweist.

35 27. Zellkultur-adaptiertes HCV-RNA-Konstrukt nach Anspruch 25 oder 26, dadurch gekennzeichnet,

daß es einen oder mehrere der in Tabelle 3 aufgeführten Nukleotid- und/oder Aminosäureaustausche aufweist, wobei Tabelle 3 Bestandteil dieses Anspruchs ist.

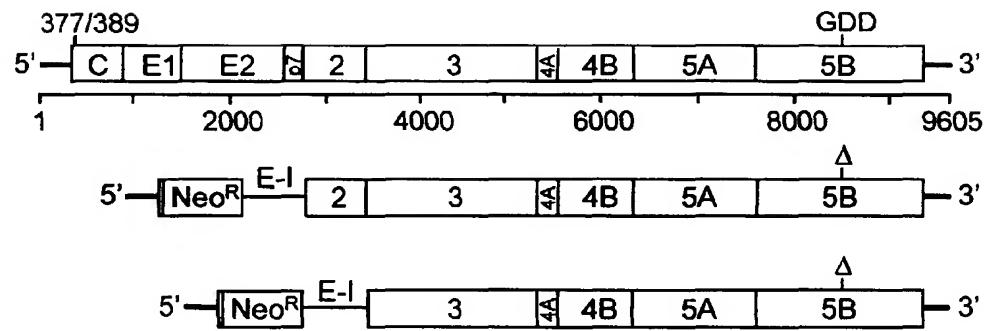
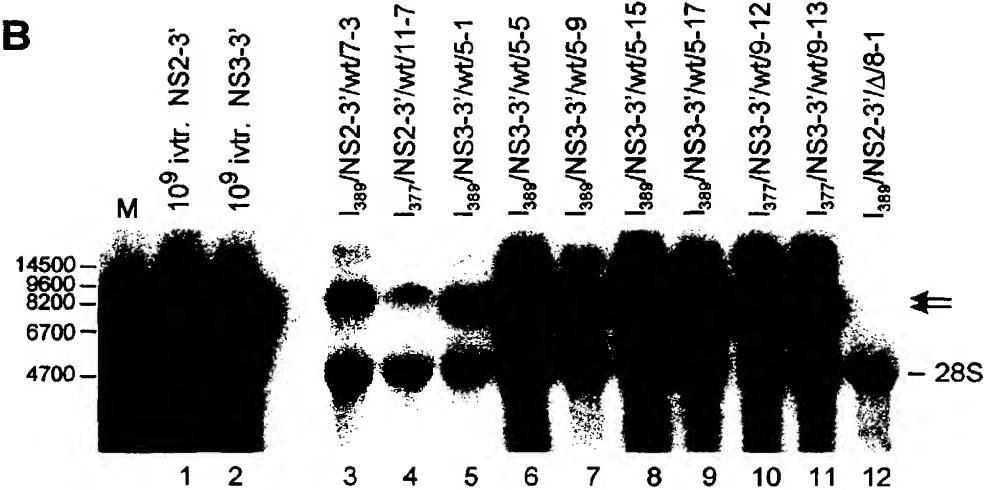
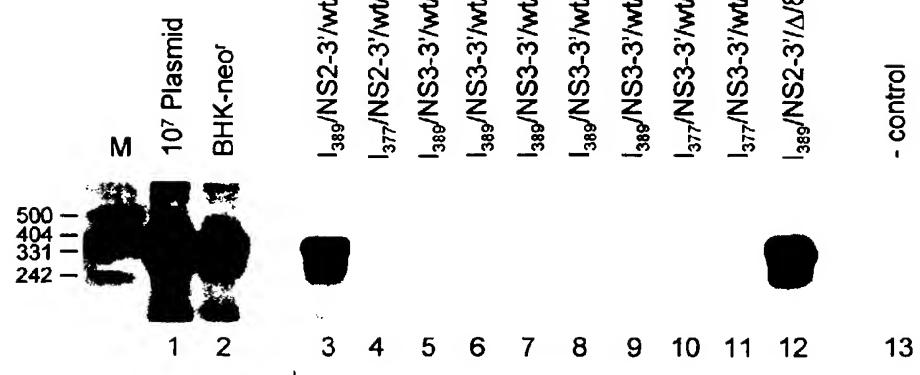
40 28. Zellkultur-adaptierte Mutanten eines HCV-RNA-Konstrukts oder eines HCV-Vollängengenoms mit im Vergleich zu dem ursprünglichen HCV-RNA-Konstrukt oder dem ursprünglichen HCV-Vollängengenom erhöhter Replikationseffizienz, dadurch gekennzeichnet,

45 daß sie mit einem Verfahren erhältlich ist, bei dem man in einem zellkultur-adaptierten HCV-RNA-Konstrukt nach Anspruch 24 durch Sequenzanalyse und Sequenzvergleich die Art und Anzahl der Mutationen bestimmt und diese Mutationen in ein HCV-RNA-Konstrukt, insbesondere in ein HCV-RNA-Konstrukt gemäß einem der Ansprüche 4 bis 15, oder in ein (isoliertes) HCV-RNA-Vollängengenom einführt, entweder durch gezielte Mutagenese oder durch Austausch von Sequenzabschnitten, die die betreffenden Mutationen enthalten.

50 29. Hepatitis C Viruspartikel oder virus-ähnliche Partikel dadurch gekennzeichnet,

daß sie mit einem Verfahren nach einem der Ansprüche 22-24 erhältlich sind.

55 30. Zellen, infiziert mit Hepatitis C Viruspartikeln oder virus-ähnliche Partikeln gemäß Anspruch 29.

A**B****C****Fig. 1**

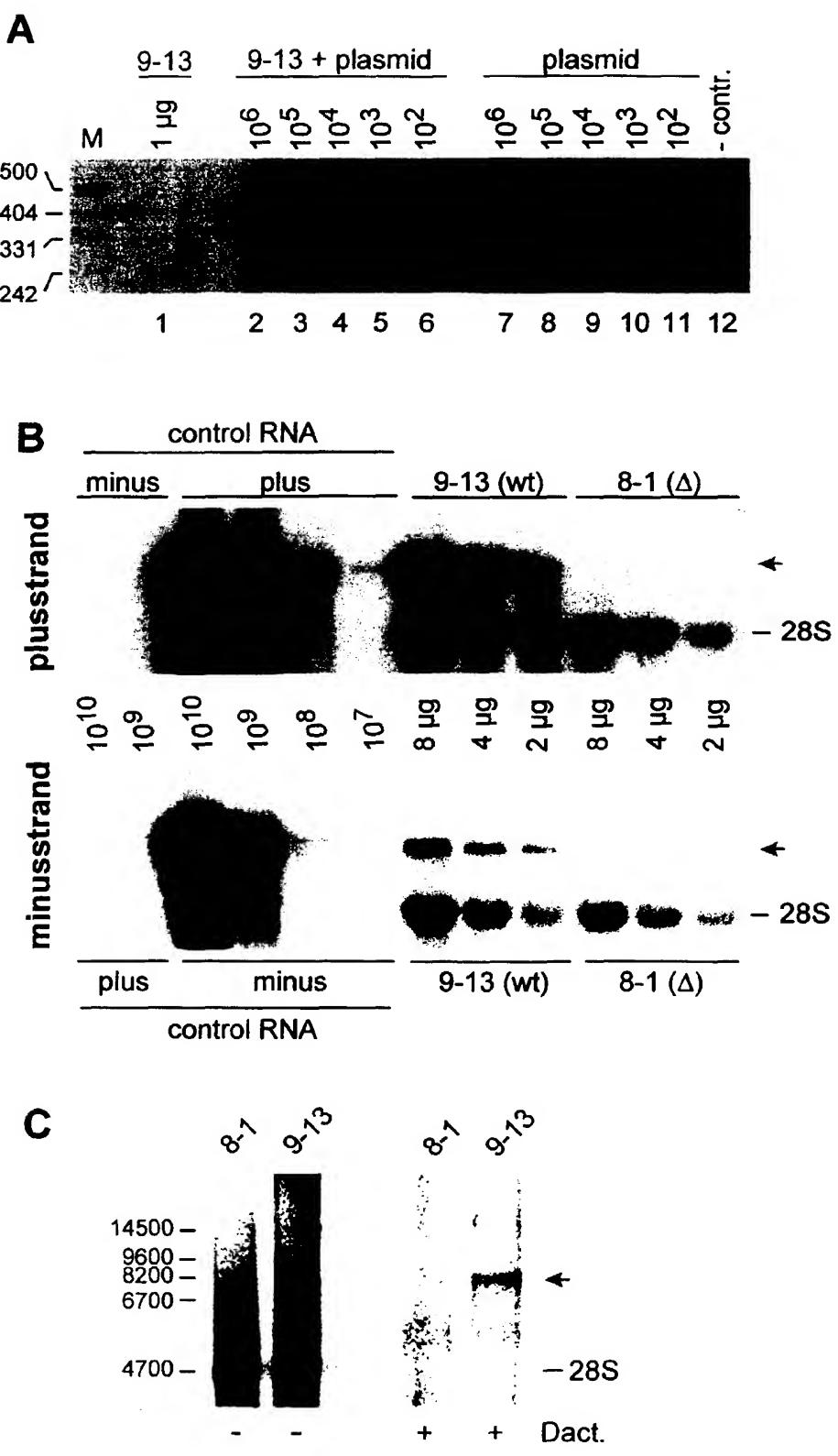
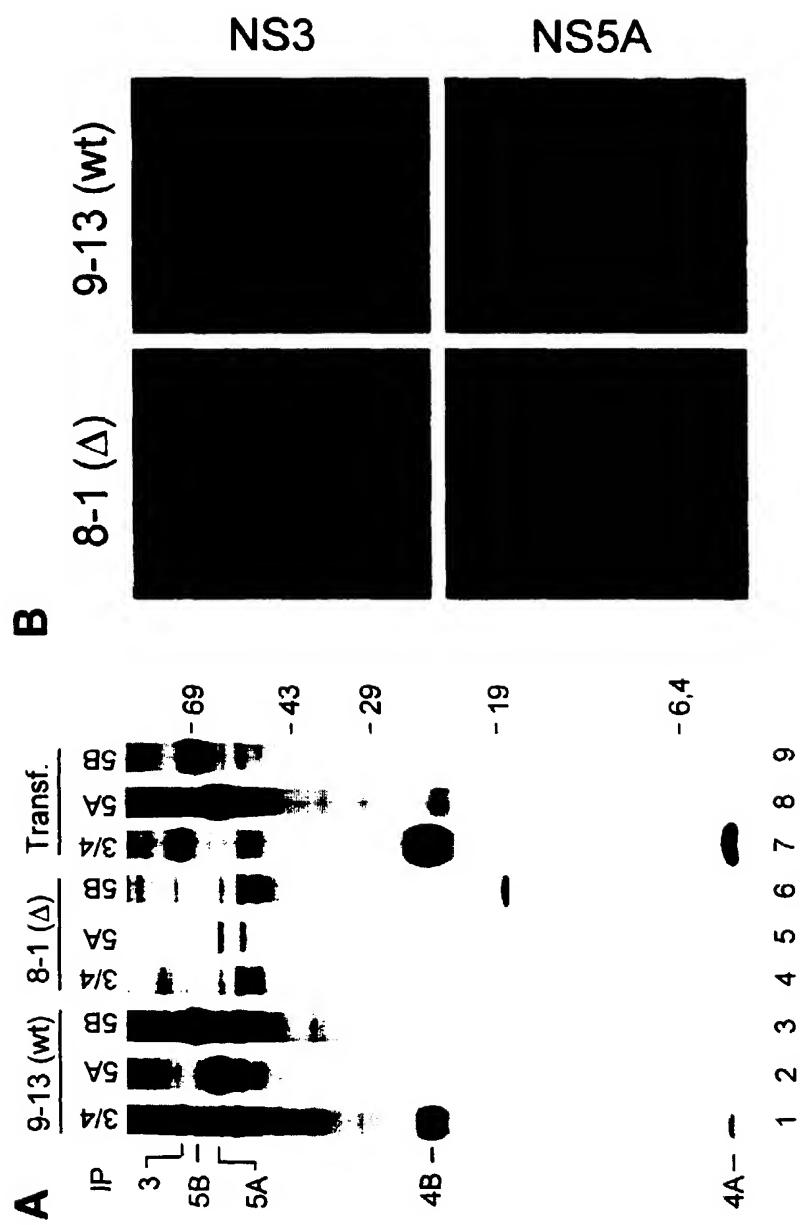


Fig. 2

Fig. 3



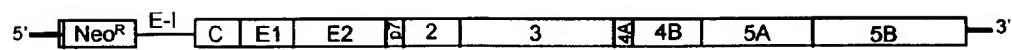


Fig. 4

A



B

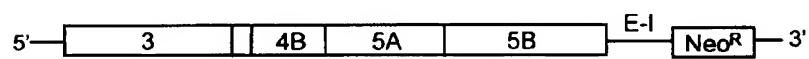


Fig. 5

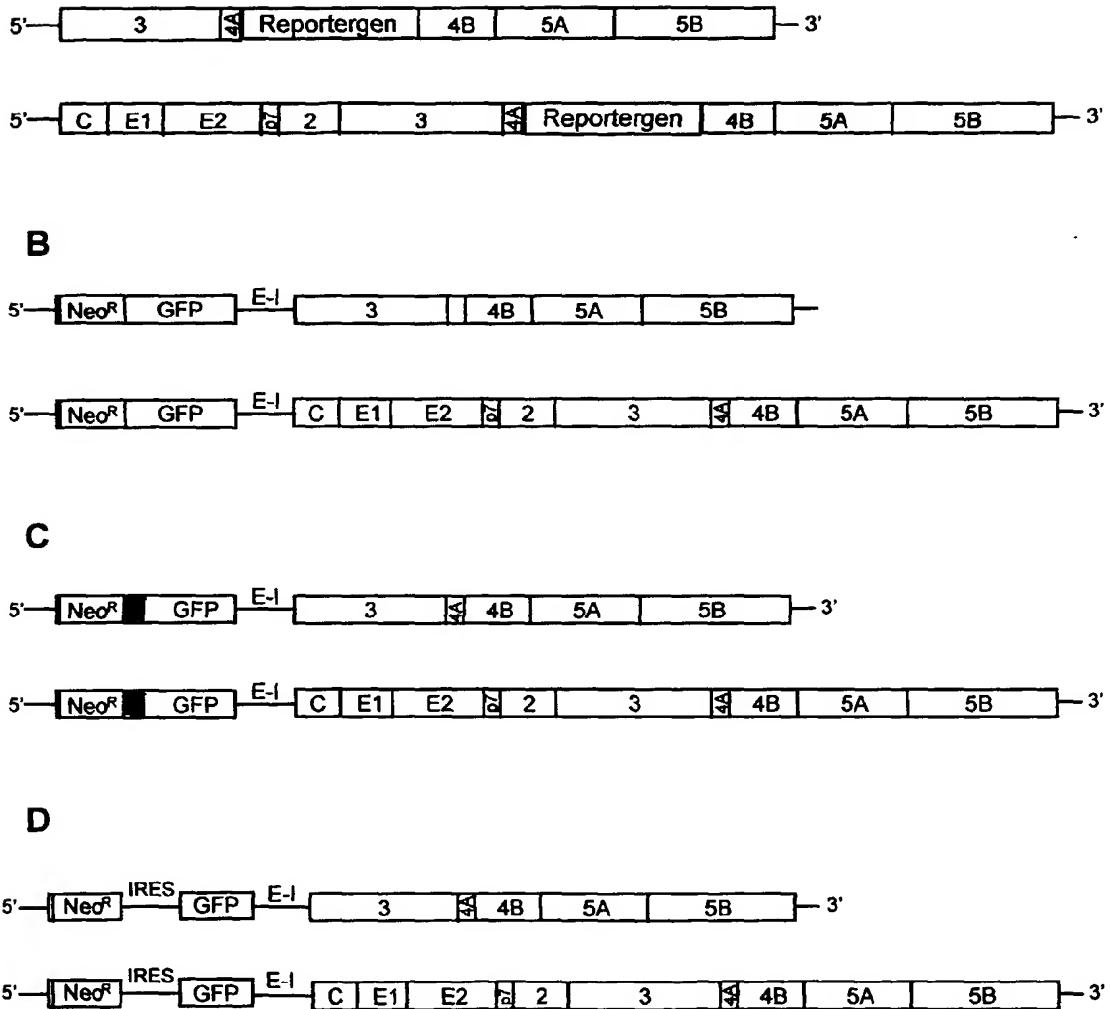


Fig. 6

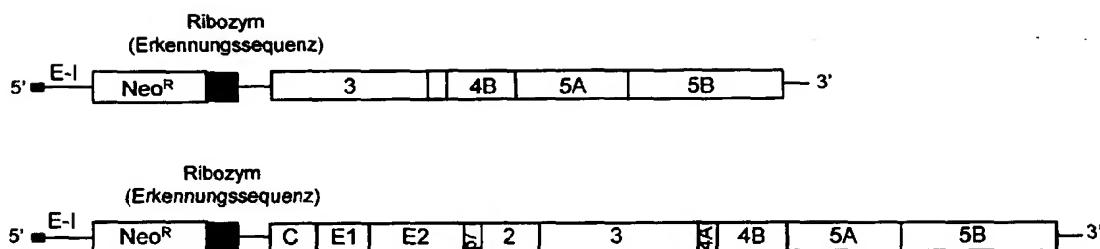


Fig. 7

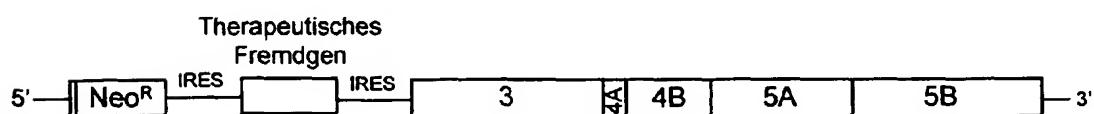


Fig. 8

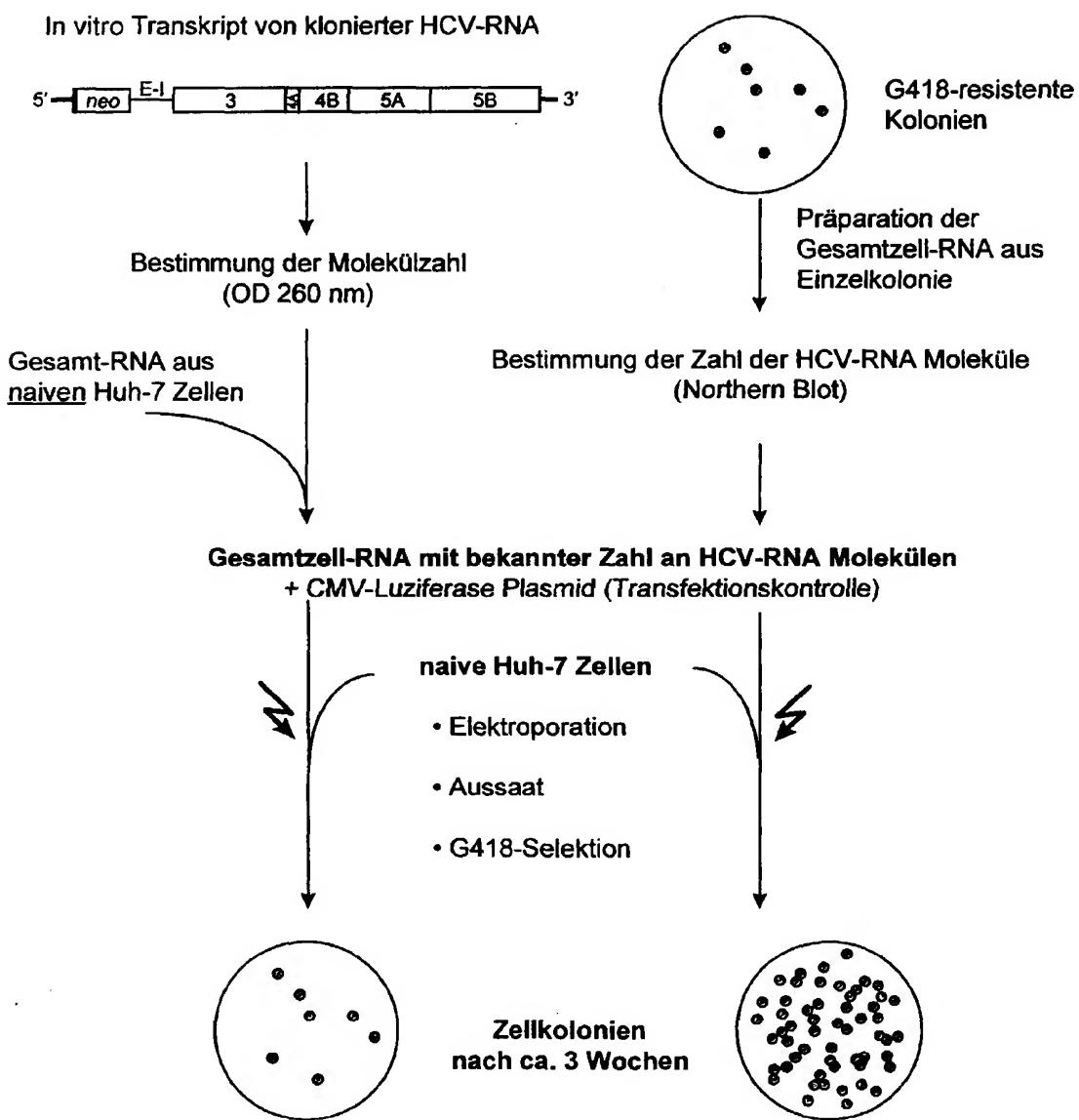
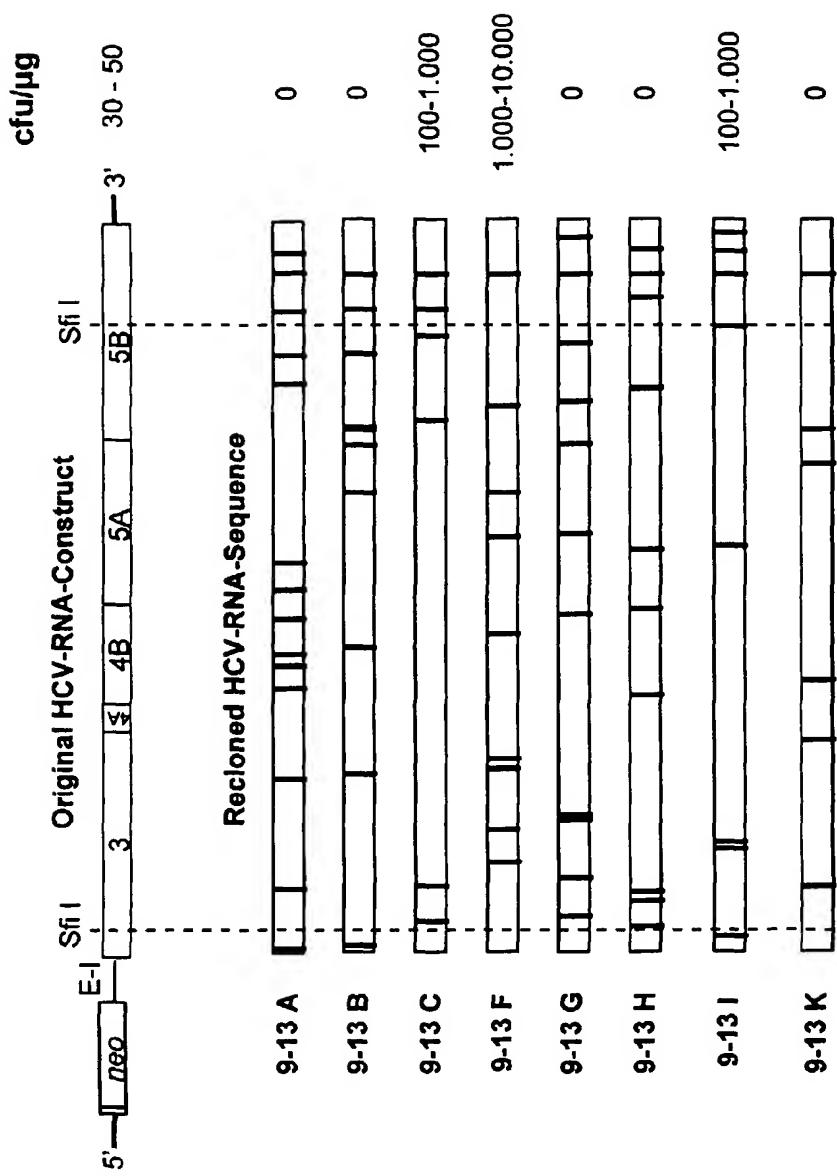
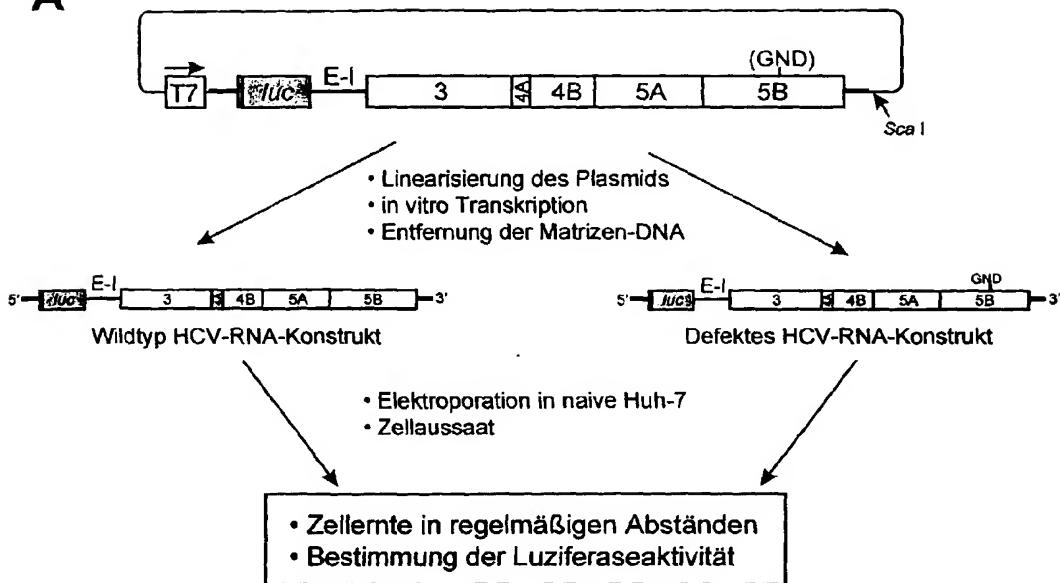
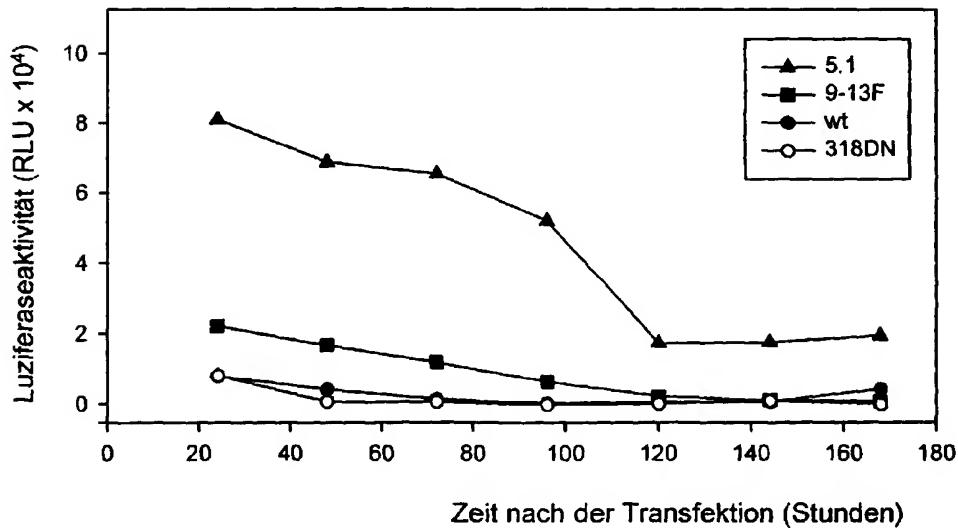
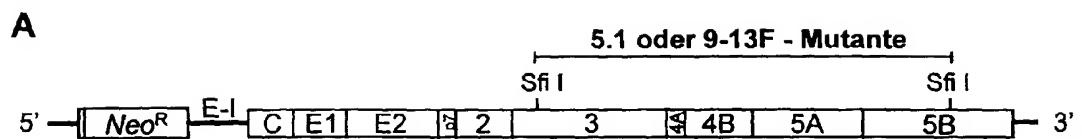


Fig. 9

**Fig. 10**

A**B****Fig. 11**



G418-Konzentration:	500 µg/ml	250 µg/ml	100 µg/ml
Anzahl erhaltener Kolonien:	1	23	75

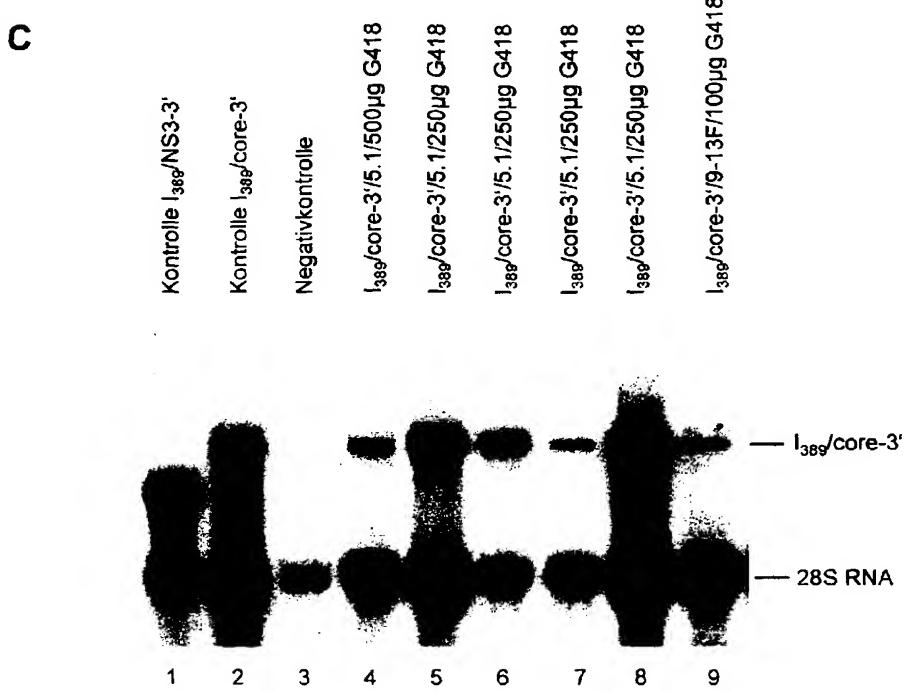


Fig. 12

Fig. 13

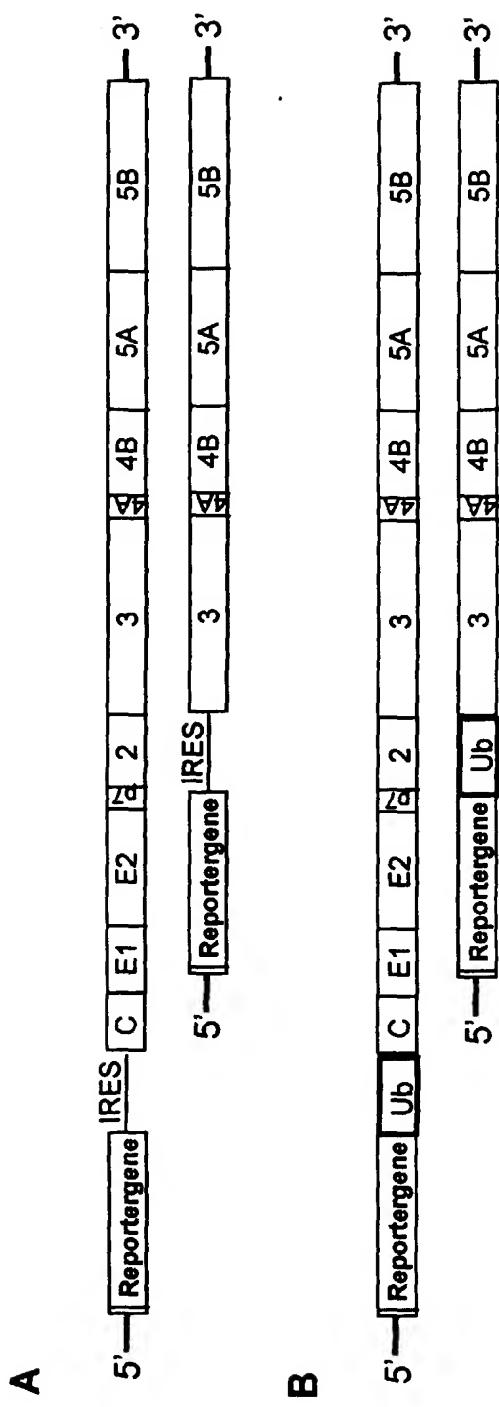


Fig. 14



Fig. 15

